460-A-3

ED 1.310/2:279222

DEC 05 1987

ED279222

COMPLETED
136

Technical Report No. 25

Error Profile: Fall Enrollment Survey of Higher Education General Information Survey

Darlene F. Russ-Eft
David A. Brandt

Prepared by

SAGE
STATISTICAL ANALYSIS GROUP IN EDUCATION

For the

National Center for Education Statistics

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

AIR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

American Institutes for Research

Box 1113, Palo Alto, California 94302

HE 020 017

BEST COPY AVAILABLE

ORIGINAL

DOCUMENT RESUME

ED 279 222 HE 020 017

AUTHOR Russ-Eft, Darlene F.; Brandt, David A.

TITLE Error Profile of Selected Statistics in the Fall Enrollment Survey of the Higher Education General Information Survey (HEGIS). Technical Report No. 25.

INSTITUTION American Institutes for Research in the Behavioral Sciences, Palo Alto, Calif.

SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.

PUB DATE Dec 82

CONTRACT 300-80-0822

NOTE 143p.; Statistical Analysis Group in Education Report.

PUB TYPE Reports - Descriptive (141) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC06 Plus Postage.

DESCRIPTORS Data Analysis; *Data Collection; *Enrollment Trends; Error Patterns; Higher Education; *Measurement Techniques; *National Surveys; Quality Control; Questionnaires; *Research Problems; *Sampling; Statistical Analysis

IDENTIFIERS *Higher Education General Information Survey

ABSTRACT

Error profiles for the Fall Enrollment Survey of the Higher Education General Information Survey (HEGIS) were developed as part of an assessment of the quality of survey data. Three statistics were of particular interest: the count of full-time equivalent students, the breakdown by race/ethnicity, and the count of unclassified students. Attention is focused on the sources of error in these three variables. To provide documentation on the sources and extent of sampling and nonsampling error, each of six stages of the survey were reviewed. The processes for the stage are described, along with possible errors arising from the process. The following stages were examined: survey objectives, sampling, measurement instruments, data collection, data preparation, data analysis and interpretation. Appendices provide background information on HEGIS, a list of nationally recognized accrediting agencies and associations, suggestions for improving the wording and format of the Fall Enrollment Survey, an example of a revised HEGIS survey form (Institutional Information for the Fall Enrollment and Compliance Report). (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TECHNICAL REPORT NO. 25

ERROR PROFILE OF SELECTED STATISTICS IN THE FALL ENROLLMENT SURVEY OF THE HIGHER EDUCATION GENERAL INFORMATION SURVEY (HEGIS)

Submitted to the National Center for Education Statistics

by

Darlene Russ-Eft

and

David Brandt

Statistical Analysis Group in Education

American Institutes for Research

P. O. Box 1113

Palo Alto, California 94302

This work was done under Contract No. 300-80-0822 with the National Center for Education Statistics, Department of Education. However, the content does not necessarily reflect the position or policy of this agency, and no official endorsement should be inferred.

December 1982

Acknowledgements

We would like to express our appreciation to the many individuals who assisted us in this effort. The following staff at the National Center for Education Statistics assisted in numerous ways by providing information, encouragement, and critical reviews: Mary Bobelis, Roslyn Korb, Samuel Peng, Andrew Pepin, and Larry Suter. Paul Berg and Anne Vaughn of the Office of Civil Rights provided us with information on their use of the Fall Enrollment Survey and located and made available needed survey forms. Tom Bukowski and John Schaeffer of VSE Corporation gave us needed documentation on their work. William Clemans and Donald McLaughlin of the American Institutes for Research reviewed drafts of these materials and made critical suggestions. Tom Bakke and Marian Eaton called representatives of California colleges and universities. Tom Bakke also assisted in some of the analyses and in the typing and proofing of the manuscript. Although all these persons contributed to this effort, the authors assume full responsibility for the content.

BLANK PAGE

Table of Contents

Page

I. Introduction 1

Definition of a Process Profile 1

Overview of this Report 2

II. Survey Objectives 5

Description of the Higher Education General Information Survey (HEGIS) 5

Purpose of the Opening Fall Enrollment Survey 9

Match between the Information Needs and the Purpose of HEGIS 9

III. Sampling 13

Definition of the HEGIS Universe 13

Criteria for listing in the Education Directory, Colleges & Universities 13

Accreditation 14

Limitations of the HEGIS Universe 15

Undercoverage 15

Overcoverage 16

Misclassification due to reclassification 17

Results of a post-validation survey 17

Effects of Frame Limitations on Enrollment Statistics 19

Sample Selection for the Early Release Study 20

IV. Measurement Instruments 23

Description of the Questionnaire 23

Potential Sources of Error 31

Match between questionnaire items and information needs ... 31

Misperception of survey objectives 32

Response difficulties 32

Wording and format problems 38

Table of Contents (Cont'd)

Page

V. Data Collection 39
  Organization and Description of the Data Collection Effort 39
  Potential Sources of Error 49
    Estimation or approximation of the enrollment data 49
    Misclassification of students 53
    Multiple counts of students 57
    Omissions of students and inclusion of dropouts 57
    Operational errors 61

VI. Data Preparation 64
  Description of the Coding, Editing, and Imputation Procedures 64
    Editing steps 64
    Procedures used to contact the schools 68
    Imputation procedures 76
  Potential Sources of Error 77
    Errors in the keypunching operations 77
    Errors in the editing operations 77
  Errors Identified by Secondary Analyses 77
    Procedures 79
    Results 80
    Discussion 85
  Errors Identified by the Post-Validation Study 87

VII. Data Analysis and Interpretation 93
  Description of the Data Presentations 93
  Potential Sources of Error 94

Table of Contents (Cont'd)

Page

VIII. Conclusions and Recommendations 95

Overview 95

Sources of Error 95

Step 1: Survey objectives 95

Step 2: Sampling 96

Step 3: Measurement instruments 96

Step 4: Data collection 97

Step 5: Data preparation 97

Step 6: Data analysis and interpretation 98

IX. References 99

APPENDIX A: Background Information on HEGIS A-1

APPENDIX B: Listing of the Nationally Recognized Accrediting Agencies and Associations B-1

APPENDIX C: Suggestions for Improving the Wording and Format of the Fall Enrollment Survey C-1

Tables

Page

Table 1. Portion of Table III--Statistics of colleges and collegiate institutions in the United States, compiled from the most recent reports sent to the United States Bureau of Education (1870) 6

Table 2. Listing and Description of the Surveys included in the Higher Education General Information Survey (HEGIS) 7

Table 3. History of Major Postecondary Education Surveys distributed by the National Center for Education Statistics (NCES) 10

Table 4. Results of Survey of California Institutions Not Included in HEGIS Listings 19

Table 5. HEGIS Categories of Enrollment Data, 1965-1978 24

Table 6. Opening Fall Enrollment in Higher Education (Fall 1972) 26

Table 7. Compliance Report of Institutions of Higher Education (Fall 1972 and Fall 1974)................... 27

Table 8. Fall Enrollment and Compliance Report (Fall 1976) 30

Table 9. Percentage Distribution of Institutions by Methods of Classifying the First-Time Freshmen and Other First-Year Students 34

Table 10. Percentage Distribution of Institutions by the Method Applied to Classify Students at the Level Higher than Freshmen 35

Table 11. Percentage Distribution of Institutions by Methods of Classifying Students to the Category of Unclassified Student 36

Table 12. Percentage of Institutions that Submitted the HEGIS Enrollment Report Through the State Coordinator 42

Table 13. Titles of Institutional Representatives Who Signed HEGIS Fall Enrollment Forms 44

Tables (Cont'd)

Page

Table 14. Information Source for Student Level Classification 45

Table 15. Information Source for Determination of Full-Time versus Part-Time Status 46

Table 16. Enrollment Record-Keeping Systems by the Type and the Control of Institutions 47

Table 17. Procedures for Compiling the Fall Enrollment Data 48

Table 18. Means, Standard Deviations, and Range of Person-Hours Required for Completing Fall Enrollment Forms 50

Table 19. Response Medium Used by Coordinated States 51

Table 20. Percentage of Schools Reporting Estimation of Enrollments by Student Level 54

Table 21. Percentage of Respondents Using Different Methods of Estimating School Enrollments 55

Table 22. Percentage of Students Misclassified as Reported by Respondents 56

Table 23. Percentage of Students Who Were Counted More than Once as Reported by Institution Personnel 58

Table 24. Percentage of Students Not Included in the Report Due to Late Registration, as Reported by Institution Personnel 59

Table 25. Due Dates Acceptable for Fall Enrollment Report 60

Table 26. Frequency of Corrections on Each Line Item in the Fall Enrollment Report 62

Table 27. Percentage of Net Errors by Data-Compiling System for Selected Student Levels 63

Tables (Cont'd)

Page

Table 28. Percentage of Schools Having Math Errors 81

Table 29. Percentage of Schools Having Page 9000 Imputed 81

Table 30. Percentage of Schools Reporting Unknown Students 81

Table 31. Percentage of Schools Using an Arbitrary Method to Distribute Reported Unknown Students 82

Table 32. Percentage of Schools Using a Data-Based Method to Distribute Unknown Students 82

Table 33. Percentage of Schools That Prorated During Editing 83

Table 34. Percentage of Schools That Used an Arbitrary Method to Prorate 83

Table 35. Percentage of Schools That Used a Data-Based Method to Prorate 83

Table 36. Percentage of Schools Reporting Students in Inappropriate Categories 84

Table 37. Percentage of Schools Found to be Out of Tolerance 84

Table 38. Net Differences Between NCES' Published Data and the "Most Accurate" Data 88

Table 39. Net Differences as a Percent of the Published Data 89

Table 40. Gross Differences Between NCES' Published Data and the Reconstructed Data 91

Table 41. Gross Differences as Percentage of the Published Data ... 92

Table A-1. Purposes for Which HEGIS Data are Used by Groups Utilizing HEGIS Data A-6

Table A-2. Uses of HEGIS Data Bank Ordered Within Groups by Purpose (1=most used) A-7

Figures

Page

Figure 1. Contents of HEGIS XIII Packages 40

Figure 2. FECR Survey Schedule 41

Figure 3. Sample FECR Survey Receipt Control Log 52

Figure 4. HEGIS Processing by the Contractor 65

Figure 5. Sample FECR Edit Cover Sheet 69

Figure 6. Edit Memo for FECR Survey 71

Figure 7. HEGIS XIII FECR Master File Status 74

Figure C-1. Reproduction of Survey Report Form C-2

Figure C-2. Example of a Revised Survey Form C-9

I. Introduction

An approach to examining the quality of survey data has recently been proposed by the Subcommittee on Nonsampling Errors of the Federal Committee on Statistical Methodology. It integrates post-hoc techniques and quality measurements into an "error profile." Specifically, a survey research team constructs an error profile by examining possible sources of error arising in each phase of the survey operation. Under the auspices of the Subcommittee on Nonsampling Errors, the U.S. Bureau of the Census prepared an error profile of employment statistics as measured by the Current Population Survey (Brooks & Bailar, 1978). The Subcommittee has, furthermore, recommended that error profiles be prepared for all important federal statistical series. Such analyses can provide useful information to those who use the resulting statistics (as an indication of the limitations in the data) and to those who produce the statistics (as a guide to areas needing improvements).

The National Center for Education Statistics (NCES), as one of the major statistical/data-gathering agencies and as a member of the Subcommittee on Nonsampling Errors, realized the value of the work on error profiles. Therefore, NCES has requested that the Statistical Analysis Group in Education (SAGE) undertake the preparation of error profiles for two major recurrent statistics. The SAGE effort should serve as a model for generating profiles on the major NCES statistics.

Definition of a Process Profile

A process or error profile documents survey procedures and identifies potential sources of sampling and nonsampling error.* The profile provides information on the quality of the survey methods and of the survey data. Thus, it can serve two audiences: (1) survey managers and designers and (2) data users. For survey managers, the profile indicates which survey procedures, if any, require improvement. For users of the survey data, the results of a process or error profile are used to describe the limitations of the data produced by the survey.

---

* The term process profile is used rather than error profile because "error" is an ambiguous term. "Error" can be interpreted in a statistical sense, or it can be used to describe the lack of precision of certain actions. Interpreted in the latter sense an "error" profile can be viewed as threatening to survey managers.

A process profile differs from a post-validation study along two dimensions: quantitative precision and completeness. The emphasis in a post-validation study is on determining the accuracy or validity of the responses to the survey. Thus the post-validation study may require extensive additional data collection and replication of selected survey procedures. In contrast, the process profile has a more general focus in examining the adequacy of the entire survey process and in determining the relative contribution of various error sources to total error. It can be viewed as a documentation and evaluation effort that is primarily conducted with available data. Secondary analyses and very limited data collection efforts, if any, are undertaken.

Overview of this Report

As part of the process profile effort, the NCES and SAGE staff met to decide on the selection of the survey and of the statistics on which to focus. One of the selected surveys was the Fall Enrollment Survey conducted as part of the Higher Education General Information Survey (HEGIS). As part of this survey, three statistics are of particular interest: the count of full-time equivalent (FTE) students, the breakdown by race/ethnicity, and the count of unclassified students. The present report focuses on the sources of error in these three variables.

To provide documentation on the sources and extent of sampling and non-sampling error, this report reviews each of six stages of the survey. The processes for that stage are described followed by a discussion of the possible errors arising from the process. The following stages are examined:

- Survey Objectives
    - Description of HEGIS
    - Needs for information on higher education
    - Purpose of HEGIS
    - Purpose of the Opening Fall Enrollment Survey
    - Method for developing and elaborating the purpose of HEGIS
    - Match between the information needs and purposes of HEGIS

- Sampling
    - Definition of the HEGIS universe
    - Sample selection for the Early Release Study
    - Potential sources of error

- Measurement Instruments
    - Description of questionnaire
    - Potential sources of error

- Data Collection
    - Organization and description of the data collection effort
    - Potential sources of error

- Data Preparation
    - Description of the coding, editing, and imputation procedures
    - Potential sources of error
    - Errors identified by secondary analyses
    - Errors identified by the post-validation study

- Data Analysis and Interpretation
    - Description of data presentations
    - Potential sources of error

The final chapter provides some suggestions for improvements in the survey, based on the findings from the error profile.

BLANK PAGE

II. Survey Objectives

Description of the Higher Education General Information Survey (HEGIS)

To understand the process of developing objectives for HEGIS, one must realize the extensive history behind this survey. The original mandate of the U.S. Office of Education (USOE), dating from 1867, required the agency to report on the condition and progress of education in the United States and its territories. Since 1869-1870, USOE has collected and reported statistics on higher education, such as enrollments, student characteristics, and faculty characteristics. Table 1 provides a portion of the data from one of the early surveys. As can be seen, many data elements of current surveys were being collected over 100 years ago. The objectives of the current survey are based, in part, on information needs identified in the 1800s.

The data collection effort increased over the years. Surveys were distributed at various times throughout the year and suffered from a lack of consistency in format, terminology, definitions, and data categories. As a result, a significant reporting burden was placed on the respondents.

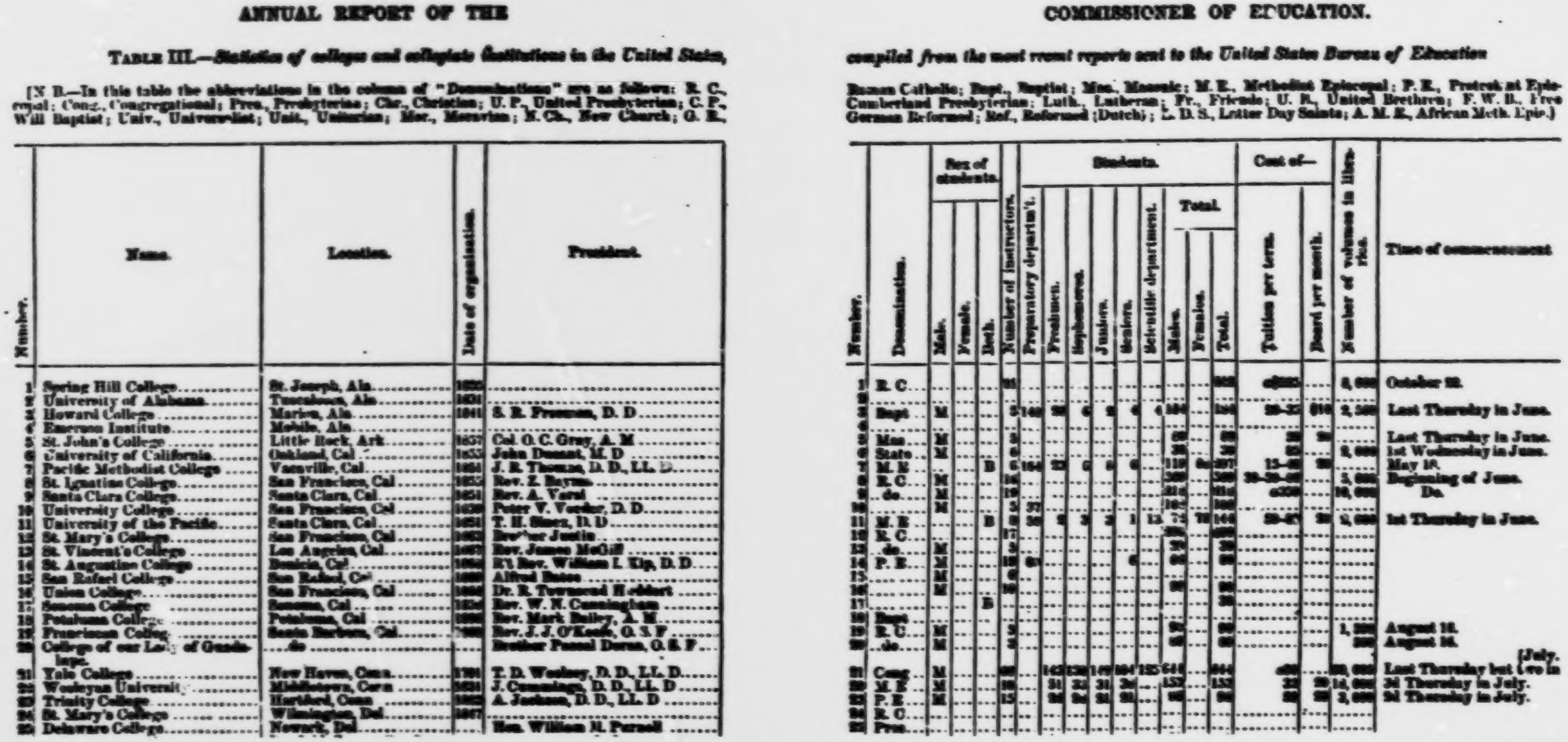
The Higher Education General Information Survey (HEGIS) was designed to overcome these problems. It assembled the old surveys, put them into a uniform format, and bundled them into a single package. This package was sent to institutions prior to the academic fiscal year to give advance notice of the data requirements, and of the data processing requirements. Furthermore, deadlines for the return of the different parts of the survey were set in accordance with times when data become available at the institutions, with most of the due dates occurring between July 15 and December 15.

The first HEGIS survey was developed for use in 1966-67, one year after the founding of NCES. It was mailed to institutions in June 1966. Although it contained some problem areas, it established the basic approach under which HEGIS has continued to function. HEGIS now consists of a package of survey questionnaires sent to over 3,000 universities and colleges. The composition of the package and the content of specific surveys varies from year to year.

The HEGIS data base consists of a complete census of public and private institutions and two-year colleges. Table 2 presents a listing and description of the surveys included in HEGIS. In addition to the active surveys listed in the table, the following surveys were once included but have been dropped from HEGIS:

Table 1

Portion of Table III--Statistics of colleges and collegiate institutions in the United States, compiled from the most recent reports sent to the United States Bureau of Education (1870)



ANNUAL REPORT OF THE COMMISSIONER OF EDUCATION.

TABLE III.—*Statistics of colleges and collegiate institutions in the United States, compiled from the most recent reports sent to the United States Bureau of Education*

[N. B.—In this table the abbreviations in the column of "Denominations" are as follows: R. C., Roman Catholic; Bapt., Baptist; Mas., Masonic; M. E., Methodist Episcopal; P. E., Protestant Episcopal; Cong., Congregational; Pres., Presbyterian; Chr., Christian; U. P., United Presbyterian; C. P., Cumberland Presbyterian; Luth., Lutheran; Fr., Friends; U. B., United Brethren; F. W. B., Free Will Baptist; Univ., Universalist; Unit., Unitarian; Mor., Moravian; N. Ch., New Church; G. R., German Reformed; Ref., Reformed (Dutch); L. D. S., Latter Day Saints; A. M. E., African Meth. Epis.]

Number.	Name.	Location.	Date of organization.	President.
1	Spring Hill College	St. Joseph, Ala	[illegible]	
2	University of Alabama	Tuscaloosa, Ala	1831	
3	Howard College	Marion, Ala	1841	S. R. Freeman, D. D
4	Emerson Institute	Mobile, Ala		
5	St. John's College	Little Rock, Ark	1857	Col. O. C. Gray, A. M
6	University of California	Oakland, Cal	1855	John Durant, M. D
7	Pacific Methodist College	Vacaville, Cal	1861	J. R. Thomas, D. D., LL. D
8	St. Ignatius College	San Francisco, Cal	[illegible]	Rev. Z. Bayma
9	Santa Clara College	Santa Clara, Cal	1851	Rev. A. Varsi
10	University College	San Francisco, Cal	[illegible]	Peter V. Veeder, D. D
11	University of the Pacific	Santa Clara, Cal	1851	T. H. Sinex, D. D
12	St. Mary's College	San Francisco, Cal	[illegible]	Brother Justin
13	St. Vincent's College	Los Angeles, Cal	[illegible]	Rev. James McGill
14	St. Augustine College	Benicia, Cal	[illegible]	Rt Rev. William I. Kip, D. D
15	San Rafael College	San Rafael, Cal	[illegible]	Alfred Bates
16	Union College	San Francisco, Cal	[illegible]	Dr. R. Townsend Hubbart
17	Sonoma College	Sonoma, Cal	[illegible]	Rev. W. N. Cunningham
18	Petaluma College	Petaluma, Cal	[illegible]	Rev. Mark Bailey, A. M
19	Franciscan College	Santa Barbara, Cal	[illegible]	Rev. J. J. O'Keefe, O. S. F
20	College of our Lady of Guadalupe	do		Brother Pascal Doran, O. S. F
21	Yale College	New Haven, Conn	1701	T. D. Woolsey, D. D., LL. D
22	Wesleyan University	Middletown, Conn	1831	J. Cummings, D. D., LL. D
23	Trinity College	Hartford, Conn	1823	A. Jackson, D. D., LL. D
24	St. Mary's College	Wilmington, Del	1867	
25	Delaware College	Newark, Del		Hon. William H. Purnell

Number.	Denomination.	Sex of students.			Number of instructors.	Students.									Cost of—		Number of volumes in libraries.	Time of commencement.
		Male.	Female.	Both.		Preparatory department.	Freshmen.	Sophomores.	Juniors.	Seniors.	Scientific department.	Total. Males.	Total. Females.	Total. Total.	Tuition per term.	Board per month.		
1	R. C.				[illegible]									[illegible]	a$[illegible]		[illegible]	October [illegible]
2																		
3	Bapt	M			5	140	[illegible]	6	[illegible]	6	4	[illegible]		[illegible]	[illegible]	$10	2,500	Last Thursday in June.
4																		
5	Mas.	M			[illegible]							[illegible]		[illegible]	[illegible]	[illegible]		Last Thursday in June.
6	State	M			[illegible]							[illegible]		[illegible]	[illegible]		2,000	1st Wednesday in June.
7	M. E.			B	6	164	[illegible]	[illegible]	[illegible]	[illegible]		119	[illegible]	[illegible]	15–[illegible]	[illegible]		May 19.
8	R. C.	M			[illegible]							[illegible]		[illegible]	[illegible]		5,000	Beginning of June.
9	do.	M			19							[illegible]		[illegible]	a$[illegible]		10,000	Do.
10		M			5	37						[illegible]		[illegible]				
11	M. E.			B	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	13	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1st Thursday in June.
12	R. C.				17							[illegible]		[illegible]				
13	do.	M			[illegible]							[illegible]		[illegible]				
14	P. E.	M			[illegible]	[illegible]				[illegible]		[illegible]		[illegible]				
15		M			[illegible]													
16		M			[illegible]							[illegible]		[illegible]				
17				B										[illegible]				
18	Bapt																	
19	R. C.	M			[illegible]							[illegible]		[illegible]			1,[illegible]	August 18.
20	do.	M			[illegible]							[illegible]		[illegible]			[illegible]	August 16.
21	Cong	M			[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	155	[illegible]		[illegible]	a[illegible]		[illegible]	Last Thursday but two in July.
22	M. E.	M			[illegible]		31	[illegible]	31	[illegible]		152		152	[illegible]	[illegible]	14,000	2d Thursday in July.
23	P. E.	M			15		[illegible]	[illegible]	21	[illegible]		[illegible]		[illegible]	[illegible]	[illegible]	2,000	2d Thursday in July.
24	R. C.																	
25	Pres.																	

BEST COPY AVAILABLE

Table 2

Listing and Description of the Surveys included in the Higher Education General Information Survey (HEGIS)

Survey	Data Items	Frequency of Collection
Institutional Characteristics	Name, address, Congressional district, county, telephone number, year established, sex of student body, previous-year enrollment, tuition and fees, control or affiliation, calendar system, highest degrees offered, type of program, accreditation, and name, title, and function of principal administrative officers.	Annual
Opening Fall Enrollment and Compliance	Full- and part-time enrollment for men and women undergraduates, graduates, first professional degree students, and unclassified students. Racial/ethnic data by selected major fields of study are collected in even-numbered years.	Annual
Earned Degrees and Other Formal Awards Conferred	First professional degrees by field; bachelor's, master's, and doctor's degrees by disciplines; and degrees and awards based on less than 4 years of work beyond high school. Racial/ethnic data for selected major fields are collected in odd-numbered years.	Annual
Residence and Migration of College Students	Student enrollment classified by residence status (in-state/out-of-state/foreign), by sex, attendance status, level of enrollment, and program of study.	Biennial (originally conducted quinquennially)
Full-Time Instructional Faculty in Higher Education	Number of faculty, by rank, sex, tenure status, and length of contract; salaries and fringe benefits of full-time faculty.	Annual
College and University Libraries	Name, address, and telephone number; number and salaries of full- and part-time staff, by sex and position; circulation and interlibrary loan transactions; book and media collections; hours and days of service; operating expenditures by source; and revenue from Federal grants.	Biennial

Table 2 (cont'd)

Survey	Data Items	Frequency of Collection
Inventory of Physical Facilities	Total gross square feet at each institution; assignable space, classified by room use and type of activity.	Biennial
Financial Statistics of Institutions of Higher Education	Current revenues, by sources (e.g., tuition and fees, government, private gifts); current expenditures, by function (e.g., institution, research, plant maintenance and operation); physical plant assets, and endowment investments and performance. Since fiscal year 1975, data are collected on changes in fund balances for each institution.	Annual
Noncredit Adult Education in Colleges and Universities	Name and address of sponsoring unit or college; field of instruction; policies regarding GED and fee remission for the elderly; registrations.	Biennial

Upper Division and Post Baccalaureate
  Enrollment (by major field)
Summer Session Enrollment
Enrollment Projections

Table 3 displays a history of the HEGIS surveys. As a total system, HEGIS provides a national data bank on enrollment and other characteristics of institutions of higher education.

Further details on the information needs and the purposes of HEGIS may be found in Appendix A.

Purpose of the Fall Enrollment Survey

The major purpose of the Fall Enrollment Survey within the HEGIS package is to collect and report on the fall enrollment count in colleges and universities in the United States and its territories. These counts are then tabulated using the following variables:

- characteristics of the institutions
  - control of institution
  - level of institution
  - state or other area

- characteristics of students
  - sex
  - attendance status
  - level of enrollment

The purpose of the Fall Enrollment Survey has expanded over the years. The following is an example. For years, NCES and the Office of Civil Rights (OCR) conducted separate but overlapping surveys. Then, in 1976, they agreed to cooperate in the conduct of a single fall enrollment survey. Such a survey would reduce the response burden on the higher education institutions by combining OCR's biennial Fall Enrollment and Compliance Report (FECR) with the Fall Enrollment Survey into one questionnaire. The agreement provides for the biennial collection of two additional data items on students: race/ethnicity and major field of study (for selected fields).

Match between the Information Needs and the Purpose of HEGIS

One potential source of error arises from the match or mismatch between the information needs and the purposes of the survey. There is a close match

Table 3

HISTORY OF MAJOR POSTSECONDARY EDUCATION SURVEYS
DISTRIBUTED BY THE NATIONAL CENTER FOR EDUCATION STATISTICS (NCES)

Actual 1966-67 through 1978-79 and Scheduled 1979-80 through 1983-84

Survey Name	Form Number	Actual													Scheduled				
		I 1966-67	II 1967-68	III 1968-69	IV 1969-70	V 1970-71	VI 1971-72	VII 1972-73	VIII 1973-74	IX 1974-75	X 1975-76	XI 1976-77	XII 1977-78	XIII 1978-79	XIV 1979-80	XV 1980-81	XVI 1981-82	XVII 1982-83	XVIII 1983-84
NCES Surveys of Institutions-- Higher Education General Information Survey (HEGIS)																			
Institutional Characteristics	2300.1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Degrees Conferred	2300.2.1	X	X	X	X	X	X	X	X	X	X	X*	X*	X	X*	X	X*	X	X*
Fall Enrollment/Compliance	2300.2.3	X	X	X	X	X	X	X	X	X	X	X*	X	X*	X	X*	X	X*	X
Residence/Migration	2300.2.8			X				X			X				X		X		X
Enrollment by Field	2300.2.9	X	X	X	X	X	X	X	X	X	X	X	Discontinued						
Employees:																			
Total Employees (including Faculty)	2300.3						X	X				X							
FT Instructional Faculty	2300.3	X	X	X	X	X				X	X		X**	X**	X**	X**	X**	X**	X**
Financial Statistics	2300.4	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Libraries	2300.5	X	X	X	X		X		X		X	X	X		X		X		X
Enrollment Projections	2300.6	X	X	X	Discontinued														
Facilities	2300.7	X		X	X	X	X			X				Study†		X			
Adult/Continuing Education	2300.8			Sample			Sample					Sample		Sample		Sample		Sample	

Source: NCES

Note: This history summarizes the years that the forms with the number and name indicated were distributed. It does not necessarily mean that the forms were consistent in structure or definition.

*Racial information required. (Form will stay the same, with racial areas shaded in years racial data not required. Thus racial information can be filled in and collected at the state level if desired.) Racial data are being collected for the Office of Civil Rights (OCR) which had collected the data previously on separate forms.

(Taken from Table 1.1 in Andrew, 1980)

**NCES agreed to collect information on salaries for continuing faculty that had previously been collected by the American Association of University Professors (AAUP). NCES also agreed to publish faculty-salary information annually.

†NCES will be modifying the Inventory of College and University Facilities to collect information for the Office of Civil Rights (OCR) regarding accessibility to higher-education facilities for mobility-impaired students. A feasibility study of 700 institutions was conducted in 1978-79, and all institutions will be surveyed in 1980-81.

BEST COPY AVAILABLE

between the information needs and the purposes of HEGIS. The primary purpose of acquiring and disseminating meaningful statistics on higher education meets the needs of numerous federal and state agencies, associations of higher education institutions, and research organizations. Evidence for this statement comes from the report by Andrew (1980) titled Analysis of uses of HEGIS data. The report examined the questions of who uses the HEGIS data and for what purpose. It summarized the results of the following activities: (1) a conventional literature review to determine the trends in HEGIS usage, (2) a statistical sampling of the relevant literature to determine the level of use, and (3) interviews and surveys of users to obtain information on the use of the HEGIS data and on opinions about its use. The following conclusions resulted concerning the information needs served by HEGIS:

> HEGIS data have provided a foundation or base for the majority of reports and books that have affected public policy on higher education.
>
> Enrollment and financial data are used much more extensively than other survey data for analyzing the Condition of Higher Education, policy analysis, and for making decisions at state and local levels.
>
> Accuracy has improved.
>
> HEGIS is a system that would have to be invented if it were not already in place because of the increasing need for data in policy making and planning.
>
> The uses of HEGIS data have increased significantly in recent years, particularly in the sophistication with which they are used.
>
> (Andrew, 1980, pp. iv-xv)

This report did, however, note some problems with HEGIS that relate to the major purpose of acquiring and disseminating meaningful statistics on higher education. These are listed below.

> Timeliness of HEGIS data is seen as a major problem.
>
> HEGIS data have not been used as extensively as they might be in reporting on the condition of women and minorities in higher education because overhead or start-up costs in using HEGIS data for analysis is relatively high.

> HEGIS is not being used as fully as it might be for policy analysis, planning and evaluation by either business or university scholars.
>
> More data are wanted on student characteristics and financial aid.

The Andrew (1980) study also indicated the partial achievement of the secondary purposes of HEGIS--to reduce the response burden and to attain consistency.

> The collection of HEGIS data has had an impact on the discipline and sophistication of data collection systems at institution and state levels.
>
> The collection of HEGIS data does not impose a heavy burden on institutions since most of the data would be collected by institutions and/or states for management purposes anyway.
>
> Institutions are concerned about the uses of HEGIS for comparison purposes.
>
> There was general agreement that data are required from all of higher education because of differences among institutions and the uses to which the data are put.
>
> HEGIS data can be used for making comparisons among sectors of higher education.
>
> (Andrew, 1980, pp. iv-xv)

III. Sampling

The following chapter describes the definition of the HEGIS universe and identifies some potential sources of errors in the universe. The final section outlines the sampling procedures used for the Early Release Survey.

Definition of the HEGIS Universe

The universe for the HEGIS Fall Enrollment survey is defined as the list of schools in the annual *Education Directory, College & Universities*, minus the system offices and central offices. This publication is prepared jointly by the Division of Eligibility and Agency Evaluation (DEAE) of the Bureau of Higher and Continuing Education of the Office of Education and by the University and College Surveys and Studies Branch (UCSSB) of the Division of Postsecondary and Vocational Education Statistics of NCES.

The HEGIS universe is controlled by a master file (the Survey Control File or SCF). The SCF is maintained by NCES on magnetic tape; it is created annually and updated cyclically with information gathered by the various HEGIS surveys. The HEGIS file for each current year is based on the HEGIS files for the previous year. Thus, the HEGIS XIII for 1978-79 consisted of all active records contained on the HEGIS XII SCF, plus the additional records of newly added institutions, minus the institutions dropped or included with other institutions. The new institutions are those that have been authorized by DEAE for inclusion in the current survey.

Criteria for listing in the Education Directory, Colleges & Universities. This *Directory* lists institutions of higher education in the United States and its outlying areas (American Samoa, Canal Zone, Guam, Northern Marianas Islands, Puerto Rico, the Virgin Islands, and the Trust Territory of the Pacific Islands). Such institutions must meet the following criteria for listing in the directory:

(1) They are legally authorized to offer and are offering at least a one-year program of college-level studies leading toward a degree;

(2) They have submitted the information required for listing; and

(3) They meet one of the following criteria for listing;

  (a) The institution is accredited at the college level by an agency that has been listed as nationally recognized by the Secretary of Education;

(b) The institution holds preaccredited status at the college level with a designated nationally recognized accrediting agency;

(c) If the institution is public or nonprofit, it has qualified under the "three-institutional-certification method" established by Section 1201 (a) (5) (B) of the Higher Education Act of 1965. By this method, the Secretary of Education verifies that not fewer than three accredited college-level institutions have accepted and do accept an unaccredited institution's credits, upon transfer, as though coming from an institution accredited by a nationally recognized accrediting agency.

"College-level" means a postsecondary associate, baccalaureate, post-baccalaureate or rabbinical education program.

Institutions listed in the Directory are asked periodically to reestablish their eligibility for listing. Institutions cited in the Directory that do not presently meet these criteria, such as the few remaining institutions originally listed on the basis of "State approval," will be granted a reasonable time and opportunity to meet these criteria.

Accreditation. No single level of government nor any public or private body or agency controls or supervises educational institutions in the United States. Accreditation of these institutions has evolved as a process involving voluntary self-evaluation by a school and appraisal by a group of peers. This process operates through nationally recognized accrediting agencies and associations and certain State bodies. These agencies or associations have established educational criteria to evaluate institutions in terms of their own objectives and to ascertain whether programs of educational quality are being maintained.

The U.S. Secretary of Education is required by statute to publish a list of nationally recognized accrediting agencies and associations determined to be reliable authorities concerning the quality of education or training offered by educational institutions or programs. To carry out this function, the Secretary has established criteria for listing nationally recognized accrediting agencies and associations. The Advisory Committee on Accreditation and Institutional Eligibility assists the Secretary in determining which accrediting bodies should be listed. Accrediting bodies that achieve initial recognition are reviewed every four years.

88

Three types of accrediting agencies currently exist. Six Regional Accrediting Commissions exist to accredit institutions with liberal arts and general programs. In some cases, special programs may be accredited. One nationally recognized State Agency, the New York State Board of Regents, recognizes degree-granting programs and curricula offered by institutions of higher education. Finally, the National Institutional and Specialized Accrediting Bodies review programs, departments, or schools that form only a portion of the total of postsecondary institutions.

In addition to full accreditation, the U.S. Secretary of Education recognizes preaccreditation status as granted by the nationally recognized accrediting agencies and associations. To be recognized, such agencies and associations must use criteria and procedures that are appropriately related to those used for the award of accreditation.

A list of nationally recognized accrediting and preaccrediting agencies and associations appears in Appendix B.

Limitations of the HEGIS Universe

The HEGIS universe is claimed to be a complete universe of institutions of higher education that meets the HEGIS requirements. The deficiencies in the frame are rather small, probably less than 1 percent of the universe. These deficiencies are described below.

Undercoverage. This results when eligible institutions are missed in the sampling process. Two major reasons exist for the misses.

1. Time lag between receipt of accreditation and entrance into frame. The following provides an indication of the magnitude of the additions of new institutions to the file from one year to the next (in this case, from 1979 to 1980).

Type and Control of Institutions	Number of Institutions Added	Percentage of Total 1978 Universe
Public		
Four-Year	2	.36%
Two-Year	6	.65%
Private		
Four-Year	23	1.65%
Two-Year	10	4.02%

These figures probably serve as an upper bound, indicating the maximum level of undercoverage due to the lag between accreditation and universe inclusion.

2. Refusal to cooperate. According to the NCES report Fall enrollment in higher education, 1978, "complete data were received for all but two of the 3,173 units in the survey" (p. 201). Assuming a base of 3,173 units, this would result in a nonresponse rate of only .06 percent. It should, however, be noted that the above statement does not agree with the description of the universe that appears in the tape documentation for the Fall Enrollment Survey, specifically:

   - "The master file tape contains data for 3,170 institutions from a universe of 3,173 reporting units. Three institutions are not included because they have enrollments which are not applicable to the Fall Enrollment Survey. [Note that this implies a base of 3,170 rather than 3,173.]

   - Six institutions provided total enrollments only, without breakdowns by specific year of study and racial/ethnic group.

   - Ten institutions did not provide racial/ethnic data.

   - Three institutions could not provide a breakdown by year of study for graduate students.

   - Two institutions could not break down their unclassified students."

   If one uses the figures given in the tape documentation, the rate of institutions failing to cooperate completely is .66 percent. This is still an extremely small percentage of the universe.

Overcoverage. This results when institutions are included in the universe but they should have been excluded. The reasons for the mistaken inclusions are as follows:

1. Time lag between loss of accreditation and removal from the frame. The following provides an indication of the magnitude of removals from the HEGIS frame because of loss of qualifications:

Type and Control of Institutions	Number of Institutions	Percentage of Total 1978 Universe
Public		
Four-Year	0	0
Two-Year	0	0
Private		
Four-Year	1	.07
Two-Year	3	1.06

   This probably indicates an upperbound for such an error.

2. Inclusion of vocational, postsecondary institutions. Depending on one's point of view, the inclusion of such institutions may or may not be considered as an instance of a problem of overcoverage. More

and more of this type of institution are acquiring accreditation status to offer associate degrees in fields that may be considered as vocational rather than academic. They do, nevertheless, meet the technical specifications for inclusion in the HEGIS universe. Thus, if their inclusion is deemed to be a problem, a special study may need to be undertaken to develop new requirements for the HEGIS universe and determine the affect of their application.

Misclassification due to reclassification. Not only are new institutions included in HEGIS and other institutions dropped from HEGIS, but institutions can change from one category to another. The following indicates the level of such reclassifications.

		New Classification			
		Public		Private	
		Four-year	Two-year	Four-year	Two-year
Previous Classification	Public				
	Four-year			1	1
	Two-year	1			
	Private				
	Four-year				
	Two-year			6	

These data indicate that most of the reclassifications resulted because of the transformation of private two-year institutions into private four-year institutions. If such transformations occurred after the HEGIS data had been collected, the institution would be misclassified during that year of the survey.

Results of a post-validation survey. A thorough examination of the accuracy of the HEGIS universe would involve a post-validation study of a sample of eligible institutions. The procedures would involve a modification of those undertaken by McLaughlin and Bakke (1981) in a validation check on the NCES nonpublic elementary and secondary school file.

Since the purpose of the present report is to provide a process profile rather than a post-validation study, we undertook an abbreviated version of the validation procedures. The purpose of the effort was to document instances of undercoverage or overcoverage. The entries in the Education Directory were compared with those in the Directory of California private postsecondary education institutions and public and private institutions "approved for veterans." Institutions that appeared in the California directory as degree-granting institutions but did not appear in the Education Directory were identified. These institutions appeared in the following categories in the California directory:

	Number of California Institutions Not Included in HEGIS
Private, Nationally Accredited Institutions	35
Institutions Offering Specified Degrees Approved by the Superintendent of Public Instruction	29
Institutions Authorized to Operate as Degree-Granting Institutions	141

All of the institutions in the first two categories and a 20 percent random sample of the institutions in the third category were contacted by telephone. Information was gathered as to familiarity with HEGIS, date of establishment, types of degrees offered, accreditation status, and enrollment counts.

The results of these telephone conversations appear in Table 6. This indicates that only three institutions that appear to be eligible for HEGIS were not included.* This equals about 1.1 percent of the California institutions in HEGIS.

The institutions of higher education located in California comprise eight percent of the total of institutions in HEGIS. If we assume that this sample is representative of such institutions throughout the country, then we can estimate that approximately 36 eligible institutions may be missed by HEGIS each year.

The opposite situation was also investigated. Four institutions appeared on the HEGIS listing that did not appear in the California directory. The telephone conversations with the institutional representatives were unable to resolve the reason for the omission of these schools from the California directory.

---

* All three institutions were private schools. Two of the three were two-year institutions and the third offered graduate degrees only.

Table 4

Results of Survey of California Institutions Not Included in HEGIS Listings

Classification of Results	Number of Institutions	Percentage of Institutions
Appears to satisfy HEGIS criteria	3	2.42
Submits information as part of another institution or branch		
Within California	16	12.90
Outside California	11	8.87
Appears to be ineligible for HEGIS		
Offers only partial credit toward degree	4	3.23
Offers only noncredit courses	13	10.48
Not nationally accredited	77	62.10
Unable to classify		
Refuses to indicate accreditation	7	—
Unable to contact (i.e., no telephone number)	74	—
TOTAL NUMBER OF SCHOOLS IN SURVEY	205	100.00

Effects of Frame Limitations on Enrollment Statistics

The limitations on the frame discussed above involve a small number of institutions, approximately 1 percent of the total universe. In addition, the problems of undercoverage and overcoverage tend to cancel each other in terms of total enrollments. This is not necessarily true for the categories of enrollment, such as grade level or ethnicity. New schools may have certain characteristics and wrongly identified schools may have other characteristics. The two factors of undercoverage and overcoverage alone seem to indicate that the effect on the total enrollment figures would be minimal, but their effect on certain categories of enrollment may be substantial.

Another factor, that of the enrollment counts of institutions included in these frame problems, also indicates that the effect on the HEGIS enrollment counts is small. The total enrollment of the three California schools omitted from HEGIS but which appear to satisfy the HEGIS criteria equals 690 or an average of 230 students per school. If we assume these figures for the entire universe, the enrollment counts may be missing as many as 8,280 students (or .07 percent of the total 1978 enrollment counts).

Sample Selection for the Early Release Study

In response to concerns about the timeliness of HEGIS data in general and of the Fall Enrollment results in particular, an Early Release Survey was developed to obtain preliminary estimates on fall enrollments in colleges and universities. Although the process profile effort concentrated on the universe survey, we will, in this section, provide a brief description of the sampling method employed for the Early Release Survey.

A two-part sampling model was used to estimate enrollments for the fall 1980 survey. The first part focused on obtaining enrollment estimates for the new institutions to be added to the HEGIS universe in 1980. A count of all the new institutions (including those who met the requirements for inclusion in 1980 but who had not done so in 1979) was compiled. Enrollments were estimated for these institutions by multiplying the number of new institutions by the average enrollment in new institutions with the same control and type. These average enrollments for new institutions were obtained from the 1976, 1977, and 1978 data.

The second part focused on the 3,190 institutions included in the 1979 universe. Enrollment data were solicited from a stratified random sample of 999 institutions. The schools to be included in the survey were chosen by the method given by Lahiri (1950; see also Cochran, 1977). This is a method of sampling proportional to size which produces unbiased ratio estimators. The universe was first stratified by type (two-year vs. four-year) and by control (public vs. private). Within each stratum the sample of size N was drawn in the following manner:

(1) Define T as the sum of the N largest values of $\underline{X}$ (total enrollment).

(2) Draw a new simple random sample of size N.

(3) Draw a new random number between 1 and T.

(4) If the sum of the N values in the sample is greater than the random number then retain the sample; otherwise reject it and return to step (2).

The estimates from the two parts of the model were added together. This provided national estimates for the fall 1980 enrollments. It should be noted that this method, unlike standard methods for probability sampling, allows unbiased estimates of variance.

BLANK PAGE

IV. Measurement Instruments

This chapter discusses the development of the Fall Enrollment questionnaire and describes the current instrument. Potential sources of errors arising from the content, wording, and format of the questionnaire are discussed.

Description of the Questionnaire

A HEGIS Fall Enrollment questionnaire has been in use since 1966. The questionnaire has been used to obtain information on student enrollments using the following dimensions:

(1) degree credit and non-degree credit
(2) resident and extension student
(3) level of enrollment
(4) full-time and part-time
(5) sex
(6) race/ethnicity
(7) type of course

Additions, deletions, and modifications have been made in the forms and in these categories across the years. Table 5 provides a summary of the categories of information gathered in each year of the survey from 1965 to 1978. This table was adapted from a similar table presented by Brown, Padgett, and Embry (1980). The table indicates that the early forms gathered limited data on resident and extension students and on total first-time degree credit students separated by sex. From 1969 to 1972, similar forms were used that asked for grade-level breakdowns for the separate categories of resident and extension. Table 6 displays the category breakdowns used in the 1972 Fall Enrollment Survey.

After 1972, it was decided to combine the enrollment counts for resident and extension students. Similar forms were used for a three-year span, from 1973 until 1975.

During this period of time, the Office of Civil Rights (OCR) gathered very similar kinds of data from higher education institutions under authorization of Section 80.6 (b) of the regulations implementing Title VI of the Civil Rights Act of 1964 (45 CDF 80.13) and similar provisions implementing Title IX of the Education Amendments of 1972. Table 7 presents the data categories included in the 1972 and 1974 Compliance Report Forms.

Table 5

HEGIS Categories of Enrollment Data, 1965-1978

	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
I. DEGREE CREDIT														
A. Resident														
1. Undergraduate														
a. Lower Division					S	S	S	S						
b. Upper Division					S	S	S	S						
Total-Undergraduate			S		S	S	S	S						
2. First-Professional					S	S	S	S						
3. Graduate			S		S	S	S	S						
4. Unclassified					S	S	S	S						
Total-Resident		S			S	S	S	S						
B. Extension														
1. Undergraduate					SP	SP	SP	SP						
2. First-Professional/ Graduate					SP	SP	SP	SP						
3. Unclassified					SP	SP	SP	SP						
Total-Extension		S	SP		SP	SP	SP	SP						
C. Resident and Extension														
1. Undergraduate														
a. Lower Division									S	S	S	X	X	X
b. Upper Division									S	S	S	Y	Y	Y
Total-Undergraduate				S					S	S	S	SCRT	SC	SCRT
2. First-Professional									S	S	S	SCRT	SC	SCRT
3. Graduate				S					S	S	S	SCRTG	SCTG	SCRTG
4. Unclassified	S	S	S	S	S	S	S	S	S	S	S	SCRTU	SCTU	SCRTU
Total-Resident and Extension														
D. First Time Degree Credit														
Total First Time Degree Credit	S	S	S	S		S	S	S	S	S	S	SCRT	SC	SCRT
II. NON-DEGREE CREDIT														
A. Resident					S	S	S	S						
B. Extension					SP	SP	SP	SP						
C. Resident and Extension	S			S	S	S	S	S	S	S	S			
D. First Time Non-Degree Credit				S		S	S	S	S	S	S			
III. GRAND TOTAL					S	S	S	S	S	S	S	S	S	S

Table 5 (cont'd)

C - Combined degree credit and non-degree credit
G - Graduate categorized by first-year and beyond first-year
P - Part-time data only
R - Classified by race: Non-resident alien, Black Non-Hispanic, American Indian or Alaskan Native, Asian or Pacific Islander, Hispanic, With Non-Hispanic
S - Data separated by sex
T - Classified by type of program: Agriculture and Natural Resources, Architecture and Environmental Design, Biological Sciences, Business and Management, Engineering, Dentistry, Medicine, Veterinary Medicine, Law, Physical Sciences, All Other
U - Unclassified categorized by undergraduate level and postbaccalaureate level
X - Lower division categorized by first-time freshmen, other first-year, and second-year
Y - Upper division categorized by third-year and fourth-year and beyond

NOTE: If space is blank, category is not available. All categories separated by full- and part-time except where noted.

Table 6

Opening Fall Enrollment in Higher Education (Fall 1972)

	HEADCOUNT					Full-time Equivalent of Total Headcount
	MEN		WOMEN		TOTAL	
	Full-time	Part-time	Full-time	Part-time		
I. DEGREE-CREDIT STUDENTS						
A. Resident Students						
1. Undergraduates						
a. Lower Division						
b. Upper Division						
c. TOTAL						
2. Unclassified Students						
3. First-Professional Students						
4. Graduate Students						
5. TOTAL						
B. Extension Students						
1. Undergraduate level						
2. Unclassified						
3. Graduate of First-Professional level						
4. TOTAL						
C. Total Resident and Extension in Degree-Credit Programs						
D. First-Time Degree-Credit Students						
II. NON-BACHELOR'S-DEGREE CREDIT STUDENTS						
A. Resident Students						
B. Extension Students						
C. TOTAL Resident and Extension Enrollment in Degree-Credit Programs						
D. First-Time Non-Bachelor's Degree Credit Students						
III. GRAND TOTAL						

Table 7

Compliance Report of Institutions of Higher Education (Fall 1972)

	(1) American Indian	(2) Negro	(3) Oriental	(4) Spanish Surnamed American	(5) All Other Students	(6) Total All Students
A. FULL-TIME STUDENTS						
1. Undergraduates						
a. First year full-time students						
b. Second year full-time students						
c. Third year full-time students						
d. Fourth and subsequent year full-time students						
e. Total number full-time undergraduate students						
2. Graduate						
a. First year full-time students						
b. Second and subsequent year full-time students						
c. Total number full-time graduate students						
3. First Professional						
a. First year full-time students						
b. Second and subsequent year full-time students						
c. Total number full-time professional students						
B. PART-TIME STUDENTS						
1. Undergraduate						
2. Graduate						
3. First Professional						

Table 7 (cont'd)
Compliance Report of Institutions of Higher Education (Fall 1974)

PART II. ENROLLMENT DATA	Black		American Indian		Asian American		Spanish Surnamed American		All Other Students		TOTAL Male	TOTAL Female	TOTAL All Students
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female			
A. Full-time Students													
1. Undergraduate													
a. First Year													
b. Second Year													
c. Third Year													
d. Fourth and Subsequent Years													
e. Total Full-time Undergraduate Students													
f. Of this total, how many are first time community/junior college transfers													
2. Unclassified Students													
3. Graduate													
a. Masters Degree													
b. Doctoral Degree													
c. Total Full-time Graduate Students													
4. First Professional Students													
B. Part-time Students													
1. Undergraduate													
2. Graduate													
3. First Professional													

BEST COPY AVAILABLE

Given the similarity in the HEGIS and OCR data requests, the *Report on the eleventh annual conference on the Higher Education General Information Survey* (1975) recommended that the two surveys be combined. The surveys were combined beginning in 1976, with the forms being the same since then. Table [illegible] presents the categories as they appear in the most recent surveys. (See Figure 1 for selected portions of the questionnaire used in 1978.) It should be noted that the combined HEGIS-OCR survey (using a detailed form) is conducted even-numbered years (e.g., 1978-79 school year), and the HEGIS Fall Enrollment Survey (using a simple card form) is conducted during the odd-numbered years. The enrollment figures are examined by OCR for undergraduate students and on a "first-year" and "beyond-the-first-year" basis for graduate students. These data enable OCR "to maintain a close scrutiny on institutions' progress in providing educational opportunities for minorities" (National Center for Education Statistics, 1978, p. 4-1). In addition to total enrollment figures, institutions must provide enrollment data by the following 10 selected fields of study: Agriculture and Natural Resources, Architecture and Environmental Design, Biological Sciences, Business and Management, Engineering, Dentistry, Medicine, Veterinary Medicine, Law, Physical Sciences, and All Other. This information allows OCR to identify discriminatory enrollment practices by selected fields of study and to determine the supply of minorities by these areas of specialization.

Another modification was made in the form as a result of a recommendation in the *Report on the eleventh annual conference on the Higher Education General Information Survey* (1975). The conference participants stated that the interpretation of the unclassified-student category is meaningless, since it includes both undergraduate and graduate students. The recommendation, which was subsequently followed by HEGIS, was to separate the unclassified student category into undergraduate unclassified students and graduate unclassified students. Thus, total undergraduate enrollment and total postbaccalaureate enrollment could be determined.

A more recent modification was included in the form based on a recommendation from a post-validation study of the HEGIS Opening Fall Enrollment Survey (Peng, 1979). The recommendation was to use a separate enrollment form for two-year colleges. "Some two-year colleges solve their enrollment classification dilemmas by lumping all students into the unclassified category. In some instances, from their point of view, this is not unrealistic,

Table 8

Fall Enrollment and Compliance Report (Fall 1976)

1. Name of Institution	2. Institution Number	3. Due Date Not Later than December 15, 1976	Form Approved OMB No. 51-R0733

9000 - Summary (Total Enrollment)		Non-Resident Alien		Black Non-Hispanic		American Indian or Alaskan Native		Asian or Pacific Islander		Hispanic		White Non-Hispanic		TOTAL (Sum of Columns (1) Through (12))		Normal Full-Time Credit Hour Load	Total Credits Enrolled For
All Students Enrolled (Resident or Extension)	Line No.	Men (1)	Wom. (2)	Men (3)	Wom. (4)	Men (5)	Wom. (6)	Men (7)	Wom. (8)	Men (9)	Wom. (10)	Men (11)	Wom. (12)	Men (13)	Wom. (14)	(15)	(16)
I. FULL-TIME STUDENTS																	
A. Undergraduates, Total																	
1. First-time freshmen																	
2. Other first-year																	
3. Second-year																	
4. Third-year																	
5. Fourth-year & beyond																	
B. Unclassified Students, Total																	
1. Undergraduate level																	
2. Post-Baccalaureate level																	
C. First Professional Students																	
D. Graduate Students, Total																	
1. First-year																	
2. Beyond the first year																	
Total Full-Time Students																	
II. PART-TIME STUDENTS																FTE of Part-Time	
A. Undergraduates, Total																	
1. First-time freshmen																	
2. Other first-year																	
3. Second-year																	
4. Third-year																	
5. Fourth-year & beyond																	
B. Unclassified Students, Total																	
1. Undergraduate level																	
2. Post-Baccalaureate level																	
C. First Professional Students																	
D. Graduate Students, Total																	
1. First-year																	
2. Beyond the first year																	
Total Part-Time Students																	
III. GRAND TOTAL ALL STUDENTS																	

~~49~~

~~48~~

BEST COPY AVAILABLE

since a student is often permitted to take any course offering without establishing his status as a full-time, part-time, undergraduate, graduate, or even high school graduate. The effort made to "force" enrollment classifications to coincide with the four-year institutions may not be the best solution" (p. 2-42). Following Peng's suggestion, two-year institutions receive a questionnaire that has many of the inappropriate categories blacked out (e.g., enrollment counts for third- and fourth-year students).

Potential Sources of Error

Errors in survey data, including nonresponse, can arise from the content, wording, and format of the measurement instrument. The following paragraphs will describe what is known about the HEGIS Fall Enrollment Survey questionnaire.

Match between questionnaire items and information needs. According to the Report on the eighth annual conference on Higher Education General Information Survey, the following is a list of the information needs concerning student characteristics:

- Educational major
- Sex
- Ethnic background
- Age
- Economic background
- Employment background
- Educational background
- Ability
- Motivation
- Attitudes
- Values
- Educational objectives

Of these information needs, the HEGIS Fall Enrollment Survey provides data on institutional enrollments by sex, ethnic background, and educational major. Studies of individual students such as the High School and Beyond Survey and the Recent College Graduate Survey do address some of the other information needs on student characteristics. Since it has been an important objective of HEGIS to keep the response burden on the institutions to a minimum, additional response dimensions have not been included in the Fall Enrollment Survey. If these information needs are, in fact, considered critical, and if items can be developed for these needs (which HEGIS respondents can answer), additional surveys might be undertaken using a sample of institutions (as is done with the Recent College Graduate Survey).

The need for information on college majors is only partially being served by the Fall Enrollment Survey. Data are gathered on enrollment in 10 selected fields to provide information on selective discrimination. Using the full range of educational majors would provide more useful information; however, including all of these categories in the questionnaire would significantly increase the response burden. This could result in a lowered response rate or a reduced reliability in the responses. As an alternative solution, certain other categories may be substituted for the current set (Berg, personal communication, 1982).

Misperception of survey objectives. Peng (1979) undertook the first post-validation study of the HEGIS Fall Enrollment Survey focusing on the fall 1977 survey. This study included interviews of a sample of 120 institutions to obtain estimates of reporting errors and reports of difficulties encountered in completing the HEGIS forms. One of the factors identified as contributing to error was labelled "inconsistent census dates and data coverage." The problem was that the census date of the survey varied across institutions; some submitted their reports to NCES promptly, while others delayed their submissions. The issue really involved a lack of information on the survey objectives and the actual information needs. Many institutions were uncertain as to the focus of the survey--the peak enrollment, the total number of students ever enrolled, the number of students currently enrolled, the students who completed the fall term, or the total number of students after the drop/add period. Thus, some institutions (56%) deleted dropouts from their reports while others (44%) included them. In addition, "many institutions reported, as they were supposed to, the enrollment data at a given census date, while others included those students enrolled after the report due date." It might be appropriate to attempt to gain some consensus among participating institutions and to include the agreed-upon objectives in the data collection instruments. This might follow the model used by the National Center for Higher Education Management Systems (NCHEMS) in the examination of portions of the Information Exchange Procedures (IEP) for their applicability and validity to the situation of major research universities.

Response difficulties. Peng (1979) also identified areas causing response difficulties. For example, of the 119 institutions that participated, 25 percent reported that they did not follow the questionnaire's definition for student levels. This problem was further examined for each major student level.

As shown in Table 9, almost 21 percent of the institutions with undergraduate students reported difficulties in distinguishing "first-time freshmen" and "other first-year students," and 14 percent use one of the following procedures rather than undertaking extensive efforts to verify the student status:

- arbitrary division between first-time freshmen and other first-year students (9.8%)
- automatic assignment to first-time freshmen (3.4%)
- excluding from the HEGIS reports those students who could not be classified (1.2%)

The results displayed in Table 10 indicated that 15% of the institutions reported difficulties in classifying undergraduate students at levels higher than freshmen. Major problems arose with the classification of transfer students. These were resolved by one of the following methods:

- placement into one category (10.1%)
- random classification (2.7%)
- other (e.g., arbitrary classification) (1.5%)

Only .7 percent consulted with department heads to determine the proper classification.

Table 11 indicates that about 15 percent of the institutions reported difficulties in responding to the category of unclassified student. Many institutions resolved this problem by combining the unclassified students into one category without a detailed breakdown of undergraduate and post-baccalaureate levels. In addition Peng (1979) mentioned a further problem with the category of unclassified students. "As reported by many institutions, this category frequently became a 'dumping' spot for placing any student for whom sufficient information was lacking at the time the HEGIS report was being prepared."

About 13 percent of the institutions reported that they did not follow the definition for indicating full-time versus part-time status. Institutions tended to over-report the full-time students and to under-report the part-time students. However, the overall effect on data accuracy was small, with the error being less than one percent.

Table 9

Percentage Distribution of Institutions by Methods of Classifying the First-Time Freshmen and Other First-Year Students

Methods	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Reported having difficulties	20.6	19.7	35.9	10.8	28.0
Assign students to the first-time freshmen	3.4	6.1	4.8	2.3	0.0
Arbitrary division	9.8	9.1	17.7	3.0	23.2
Extensive efforts to verify student status	6.2	4.5	9.3	5.5	4.8
Excluding students who cannot be classified	1.2	0.0	4.2	0.0	0.0
Reported having no problem with NCES' definitions	79.4	80.3	64.1	89.2	72.0
Sample Size	119	26	35	43	15

(Taken from Table 2.1 in Peng, 1979)

Table 10

Percentage Distribution of Institutions by the Method Applied to Classify Students at the Level Higher Than Freshmen

Methods	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Reported having difficulties:	15.0	22.6	24.2	9.2	4.8
Randomly classify transfer students	2.7	8.0	1.4	2.0	0.0
Put all students into one category	10.1	12.8	18.6	4.7	4.8
Consult with department heads	0.7	1.8	1.4	0.0	0.0
Other (e.g., arbitrary classification	1.5	0.0	2.9	1.4	0.0
Reported having no problem with NCES' definitions	85.0	77.4	75.8	91.8	95.2
Sample Size	119	26	35	43	15

(Taken from Table 2.2 in Peng, 1979)

Table 11

Percentage Distribution of Institutions by Methods of Classifying Students to the Category of Unclassified Student

Methods	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Reported having difficulties:	14.3	26.6	19.2	8.3	8.5
Reported one category with no breakdown of undergraduate and post-baccalaureate levels	8.9	15.0	14.7	2.9	8.5
Arbitrary assignment	3.6	9.7	1.8	3.3	0.0
Had special program	2.1	2.0	2.7	2.1	0.0
Reported having no problem with NCES' definition	85.4	73.4	80.8	91.7	91.5
Sample Size	119	26	35	43	15

(Taken from data reported in Peng, 1979)

Another problem area consists of the definition of full-time equivalent (FTE). The following are the instructions given for computing the FTE enrollment of part-time students.

(1) Use a method already employed in your institution to compute FTE's for some other purpose.

(2) Sum the credit hours for part-time students and divide by the normal full-time credit-hour load. (NORMAL FULL-TIME CREDIT-HOUR LOAD) is usually determined by dividing the total number of credits required for completing the program by the number of terms normally required to obtain them. Do not confuse this with the minimum number of credit-hours required for a student to be classified full-time (75% of a normal full-time load). NOTE: Divide by the normal, or average, full-time load, not by the minimum full-time load. For most institutions, this will be 15 credit-hours (not 12).

(3) Assign a fractional value of full-time to each part-time student, appropriate to your institution, such as 1/4, 1/3, or 1/2. Remember that a student taking 3/4 (75%) or more of a normal full-time load should be classified as a full-time student.

The latter two methods require some calculation using credit-hour load or a proxy (e.g., tuition fees based on credit-hour load). According to the Peng (1979) study 84 percent of the institutions used credit-hours and 1 percent used tuition fees as the information source for classifying students as full-time or part-time. The remaining 15 percent used some other method. From this, we can assume that at least 15 percent of the institutions are using their own method for calculating FTE (option 1). This makes examination of the FTE data for potential error difficult, if not impossible. One alternative would involve a post-validation study of the methods used by institutions to calculate FTE.

Another problem arises from the provision for alternative methods in calculating FTE. By allowing institutions to choose the method for calculating FTE, unknown inconsistencies arise among institutions. Recognizing this problem, the report of the student workshop included in the Report on the eleventh annual conference on the Higher Education General Information Survey (1975) recommended that the form include an item requesting schools to indicate the method used for calculating FTE. This recommendation has not yet been followed.

Wording and format problems. The wording and format of a questionnaire can be critical to the gathering of accurate data. [See the SAGE Technical Report on reliability and validity of survey data by Russ-Eft (1980) for a discussion of some of the important considerations in questionnaire development.] Appendix C provides some suggestions for improving the wording and format of the HEGIS questionnaire. It should be recognized that the major burden on the respondents is the retrieving and compiling of the information. Nevertheless, reducing the burden in reading the survey form may improve response accuracy.

V. Data Collection

This chapter describes the data collection procedures used in the Fall Enrollment Survey. The procedures are completed by several different parties: NCES, the survey contractor (i.e., VSE Corporation), state coordinators, and institutions. Errors may occur at each stage. Thus, this chapter will identify the procedural contributions to survey error.

Organization and Description of the Data Collection Effort

The HEGIS package is mailed each year to all institutions included in the previous year's Education Directory plus any newly eligible institutions. Thus the HEGIS XIII package was mailed in early October based on the HEGIS XII Education Directory. Figure 1 displays the contents of the entire HEGIS package. Figure 2 presents the schedule for the data collection and processing operations for the Fall Enrollment survey. These figures are taken from the HEGIS XII Requirements and Specifications Manual (1978).

The questionnaire forms are sent to the institutions or to a state coordinator, depending on the type of institution and its location. In 37 states, the materials were mailed to the state coordinator, who then distributed them to the individual institutions. In seven states, materials were mailed to state agencies for some of the schools (e.g., public institutions), while materials were mailed directly to other institutions within the state (e.g., private schools). In the remaining states, materials were mailed directly to the institutions. (Further details on the mailing procedures may be found in the HEGIS XIII Requirements and Specifications Manual.) Table 12 from the Peng (1979) post-validation study indicates that a state coordinator was involved in the data collection process for a majority of the schools. In some cases, the states actually prepared the HEGIS report based on data collected from individual institutions.

In addition, most institutions, even those without a state coordinator, were required to prepare an enrollment report for their state. Results from the post-validation study indicated that 89 percent of the public four-year institutions, 94 percent of the public two-year institutions, 62 percent of the private four-year institutions, and 89 percent of the private two-year institutions prepared enrollment reports for the state. Among the institutions who prepared separate state reports, 64 percent indicated that these

Figure 1

CONTENTS OF HEGIS XIII PACKAGES

Item	Quantity
Letter: to the President[1]	1
General instructions for the completion of all survey forms (to each of which is affixed an undesignated label)	3
Instructions for updating the printout for the *Directory of Colleges and Universities, 1978-79*	1
Return labels (addressed to NCES)	7
Basic Forms (to each of which is affixed a "FROM" label): 2300-2.3 Fall Enrollment and Compliance Report[2] 2300-3 Salaries, Tenure, and Fringe Benefits 2300-4 Financial Statistics	3
Additional forms (to each of which is affixed a from label): 2300-2.1 Degrees and Other Formal Awards 2300-1 Institutional Characteristics[3] 2300-8 Adult/Continuing Education[4]	2-3

[1] Only one letter is sent to each institution.

[2] Letter from the Office for Civil Rights is included.

[3] Previously eligible institutions receive two copies of a computer printout displaying data reported the previous year. Newly eligible institutions receive two copies of NCES Form 2300-1 (with no "FROM" label affixed). Multicampus institution packages are clamped together.

[4] If the institution is included in the sample universe, one copy of a special letter is included.

(Taken from the HEGIS XIII Requirements and Specifications Manual)

Figure 2

FECR SURVEY SCHEDULE

EVENT	DATE FOR PRELIM RPT	DATE FOR FINAL RPT
Mail-Out of questionnaire	Oct. 3, 1978	Oct. 3, 1978
Edit programming initiated	Oct. 10, 1978	Oct. 10, 1978
Output programming initiated	Nov. 1, 1978	Nov. 1, 1978
Receipt Control Log activated	Oct. 6, 1978	Oct. 10, 1978
Reminder Letter mailed	N/A	Oct. 30, 1978
Due date of completed questionnaire	Oct. 15, 1978	Nov. 15, 1978
Premachine edit initiated	Oct. 6, 1978	Oct. 10, 1978
Premachine edit followup initiated	Oct. 16, 1978	Oct. 10, 1978
First followup letter mailed	N/A	Nov. 27, 1978
Second followup (mailgram) mailed	N/A	Dec. 5, 1978
Followup telephone calls initiated	Oct. 16, 1978	Dec. 13, 1978
Machine edit program operational	Nov. 1, 1978	Nov. 1, 1978
Error resolution process initiated	N/A	Dec. 11, 1978
Programmed table shells delivered to survey director	N/A	Feb. 1, 1979
Nonresponse imputation for Preliminary Report initiated	Dec. 14, 1978	N/A
Preliminary Report closeout	Dec. 8, 1978	N/A
Clean Data Base for Preliminary Report declared	Dec. 13, 1978	N/A
Preliminary Report lined tables delivered	Dec. 15, 1978	N/A
Preliminary Report white paper tables delivered	Dec. 19, 1978	N/A
Nonresponse imputation for final report initiated	N/A	N/A
Table-generation programs operational	Dec. 11, 1978	Jan. 9, 1979
Sample tables delivered to survey director	N/A	N/A
Operational closeout	N/A	June 29, 1979
Clean Data Base declared	N/A	Aug. 24, 1979
Table-Ready Data Base achieved	N/A	Aug. 29, 1979
Lined tables edited and delivered to survey director	N/A	Sept. 19, 1979
Review of lined tables completed by survey director	N/A	Oct. 5, 1979
Delivery of white paper tables to survey director	N/A	Oct. 15, 1979

Requirements and Specifications Manual:

	Vol. I	Vol. II	Vol. III	Vol. IV
First Draft to UCSSB	Oct. 6, 1978	Nov. 14, 1978	Feb. 6, 1979	N/A
Camera Ready Copy to UCSSB	Jan. 5, 1979	Feb. 23, 1979	June 1, 1979	Jan. 5, 1979

(Taken from the HEGIS XIII Requirements and Specifications Manual)

Table 12

Percentage of Institutions that Submitted the HEGIS Enrollment Report Through the State Coordinator

	All Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Through the State Coordinator	67.6	57.3	71.4	71.1	56.5
Directly to NCES	24.3	17.8	17.5	28.9	36.1
Other (e.g., State prepared the report)	8.0	24.9	11.1	0.0	7.4
Sample Size	119	26	35	43	15

(Taken from Table 2-17 in Peng, 1979)

reports differed from HEGIS in terms of the items or the definitions, and 35 percent indicated that the reporting period differed from HEGIS.

The "respondents" at the colleges and universities vary from the president to a secretary. Table 13 displays the titles of the persons who signed the Fall Enrollment questionnaire as the "respondent." These people may or may not be the actual respondents, but they are the persons who are responsible for the report. The real question, however, is whether the responsible person had the necessary resources to obtain and report accurate counts.

The enrollment data reported by the institutions were gathered from different sources. Table 14 shows the information source used for determining the student level classification. About 70 percent of the institutions use the credit-hours registered for the fall term, but over 6 percent used the students' self-classification. Table 15 displays the information source used for determining full-time or part-time status. In this case, about 84 percent of the institutions used the credit-hours registered for the fall term.

The enrollment record-keeping systems used by the institutions varied from completely computerized to completely manual. Table 16 from the post-validation study indicates the kinds of systems used by the institutions. Almost half of the schools use a computer file for storing enrollment records. However, while over 90 percent of the public four-year institutions used computer files, less than eight percent of the private two-year institutions did so. Among the latter group, about 62 percent maintained card files.

The record-keeping system used by the institutions affected the procedures undertaken for responding to the survey. Table 17 shows the procedures used. Thus, over 61 percent of the four-year public institutions generated their enrollment reports using only the computer, while almost 29 percent used a partly computerized and partly manual operation. In contrast, only eight percent of the two-year private colleges used the computer for generating their enrollment reports, with another 16 percent using a partly computerized and partly manual operation. However, almost 73 percent compiled the fall enrollment data manually from a card file.

The record-keeping systems and data collection operations also affected the level of effort required to complete the Fall Enrollment report. Table 18 from the post-validation study indicates that the average hours to complete the forms were 16 hours for the card form and 33 hours for the detailed

Table 13

Titles of Institutional Representatives Who Signed HEGIS Fall Enrollment Forms

Titles	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Registrar/Associate Registrar	39.9	28.1	17.2	63.5	22.2
Director/Manager of Institutional Research	6.7	20.1	1.7	3.5	0.0
Dean, Vice President, Vice Rector, Assistant Dean, Director	4.1	0.0	3.4	3.5	16.7
Dean/Director/Assistant Dean of Admissions	4.1	0.0	10.3	1.2	5.6
President	3.1	0.0	1.7	2.4	16.7
Dean/Director of Student Services	3.1	0.0	8.6	0.0	1.2
Dean, Director, Vice President of Administrative Services/ Development/Financial Aid/Financial Affairs/ Records/Research & Planning/Students	9.3	15.6	10.3	7.1	5.6
Research Associate/ Research Assistant/ Records Assistant/ Data Technician/ Secretary	9.3	21.9	17.2	0.0	5.6
Assistant/Recorder to President/Vice President/ Provost/Dean	4.7	0.0	5.2	7.1	0.0
Consultant/Information Specialist/Statistician	3.1	3.1	1.7	4.7	0.0
Illegible/No title	13.0	3.1	22.4	8.2	27.7

Table 14

Information Source for Student Level Classification

Information Source	All Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Student's self-classification	6.5%	4.9%	9.5%	6.4%	0.0%
Credit hours registered for fall term	70.3	64.8	65.7	74.2	76.6
Tuition fees	0.0	0.0	0.0	0.0	0.0
Other	23.2	30.4	24.9	19.4	23.4
Sample Size	119	26	35	43	15

(Taken from data reported in Peng, 1979)

Table 15

Information Source for Determination of Full-Time versus Part-Time Status

Information Source	All Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Student's self-classification					
Credit hours registered for fall term	84.3%	85.9%	98.1%	74.2%	88.4%
Tuition fees	1.0	3.7	0.0	.8	0.0
Other	14.7	10.4	1.9	25.0	11.6
Sample Size	119	26	35	43	15

(Taken from data reported in Peng, 1979)

Table 16

Enrollment Record-Keeping Systems by the Type and the Control of Institutions

Record-Keeping System	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Card file	27.9	4.9	8.6	42.3	62.3
Computer file	49.1	90.6	75.6	24.5	7.7
Both computer file and card file	11.5	4.5	7.4	14.1	26.3
Others (directory class file)	11.5	0.0	8.3	19.1	3.7
Sample Size	119	26	35	43	15

(Taken from Table 2-12 in Peng, 1979)

Table 17

Procedures for Compiling the Fall Enrollment Data

Procedures	All Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Total	100.0%	100.0%	100.0%	100.0%	100.0%
Generated by computer	39.7	61.2	69.8	18.7	7.7
Compiled manually from a card file	30.9	8.0	6.1	47.5	72.6
Aggregated from department reports	.4	0.0	0.0	.9	0.0
Partly computerized/ partly manual	16.0	28.8	12.6	13.4	16.0
Other	12.9	1.9	11.5	19.4	3.7
Sample Size	119	26	35	43	15

(Taken from Table 2-13 in Peng, 1979)

form. The variation among institutions was substantial. This variation was probably related to the size of the student enrollment and to the operations used for compiling the report. Some institutions required about 10 minutes to transcribe the data from a computer printout to a report form, while others spent days to compile the information manually. For this latter group of institutions, the burden of responding to the survey may decrease the motivation and ability to provide accurate data.

The variety of data collection operations results in different response formats. The schools in most states, even those with a state coordinator, submitted the data on unedited NCES forms. Table 19 indicates the response formats used by states having a state coordinator.

The goal of the data collection effort is to obtain a 100 percent response. Follow-up is the process used by NCES to remind and urge the institutions to respond in a timely fashion. These activities are directed toward institutions appearing in a Receipt Control Log from which a valid response has not been received as of a certain date. An example of the Receipt Control Log for the Fall Enrollment survey can be found in Figure 3.

In the past, NCES staff logged each form when it was returned by the state coordinator or by the institution. Because of budget and staff reductions, the NCES staff no longer log the forms upon their return. Each day batches of forms are taken from NCES and delivered to the data processing contractor, VSE Corporation (VSE). Upon delivery the forms are grouped by area and are logged.

Potential Sources of Error

Several sources of error can arise from the data collection operations undertaken by the HEGIS Fall Enrollment survey. Below, we present and discuss available data on the effects of those errors. Most of these data come from the Peng (1979) post-validation study.

Estimation or approximation of the enrollment data. Interviews of the institutional personnel responsible for the Fall Enrollment reports revealed that about 13 percent of the institutions had to estimate the enrollment counts for all or several levels of students.

Table 18

Means, Standard Deviations, and Range of Person-Hours Required for Completing Fall Enrollment Forms

	All Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Card Form					
Mean	16.1	14.2	7.8	23.1	5.8
Standard Deviation	72.7	22.1	10.5	103.2	7.5
Range					
Lowest	.1	.1	.3	.2	.1
Highest	650.0	109.0	40.0	650.0	35.0
Detailed Form					
Mean	32.5	42.9	21.8	39.0	10.4
Standard Deviation	86.9	88.9	25.0	112.1	8.4
Range					
Lowest	.2	.2	.5	.5	1.0
Highest	709.0	709.0	80.0	700.0	40.0
Sample Size	119	26	35	43	15

(Taken from Table 2-15 in Peng, 1979)

Table 19

Response Medium Used by Coordinated States

All Coordinated States[1] submit that data on unedited NCES forms, with the following exceptions:

Response Medium	State or Area
Edited NCES Form	Pennsylvania - except community colleges
State Forms	California - University of California only
	Florida - public 2-year community colleges only
	Pennsylvania - community colleges only
Magnetic Tape	Maine - except University System
	Maryland
	Missouri
	New York - CUNY and SUNY only
	Oklahoma
	Vermont
	Washington - community colleges only
Printouts	Ohio - Miami University only
	Virginia - community colleges only
NCES Imputes Data	New York - rabbinical schools only

---

[1] Most U.S. Service Schools submit data on the NCES forms through the State Coordinator; the U.S. Merchant Marine Academy and the U.S. Military Academy also use the NCES forms, but submit the forms directly to NCES.

Figure 3

SAMPLE FECR SURVEY RECEIPT CONTROL LOG

```
HEGIS XIII                                      FALL ENROLLMENT AND COMPLIANCE REPORT      DATE 09/20/78      PAGE  0001
                    ALABAMA                                 RECIEPT CONTROL LOG

ALPHA                                     C T S  L   |      |      |     |--PRELIM--|BTCH|-----FINAL-----|MACH|
SEQUENCE         INSTITUTION        FICE  T Y T  V   |      | BY   |DUP/| KEY | TO  | #  |EDIT| KEY |     |EDIT|      REMARKS
CODE                NAME            CODE  L P A  L   |FINAL |PHONE |REV |PUNCH|FILE|    |COMP|PUNCH|     |CYCL|
-----------------------------------------------------------------------------------------------------------------------------
10-0010  ALABAMA A & M UNIVERSITY  001002 1 2 5 07   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0020  ALABAMA CHRISTIAN COLLEGE 001003 2 3 5 03   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0030  ALA LUTH ACAD AND COLLEGE 010554 2 3 5 03   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0040  ALABAMA STATE UNIVERSITY  001005 1 2 5 07   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0050  ALEXANDER CITY STATE JC   001007 1 3 5 03   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0060  ATHENS STATE COLLEGE      001008 1 2 5 04   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0070  AUBURN U ALL CAMPUSES     008695 1 7 2 08   | ---  | ---  | -- | --- | --  | -- | -- | --- | --  | -- |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0090  AUBURN U MAIN CAMPUS      001009 1 1 6 08   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0100  AUBURN U AT MONTGOMERY    008310 1 4 7 07   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0110  BIRMINGHAM STHN COLLEGE   001012 2 2 5 04   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0120  BREWER STATE JR COLLEGE   009134 1 3 5 03   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
10-0130  CHATTAHOOCHEE VALLEY CC   012182 1 3 5 03   |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
                                                     |      |      |    |     |     |    |    |     |     |    |
-----------------------------------------------------------------------------------------------------------------------------
```

BEST COPY AVAILABLE

Type and Control of Institutions	Percentage Reporting Estimation of Enrollment
Public Four-Year	9.8
Public Two-Year	22.1
Private Four-Year	10.9
Private Two-Year	0.0
Total	13.1

Table 20 shows the responses of these schools as to which student levels were estimated. The levels ranged from first-year freshmen to graduate students. Only the public two-year institutions indicated that they estimated freshmen (12.5%) and first-time freshmen (17.7%); however almost 12 percent of the private four-year schools stated that they estimated lower division students (including freshmen).

The schools who estimated student enrollments were also asked to describe the method that they used. Table 21 displays these results. Most schools reported using some percentage of past reports or making a simple deduction.

Misclassification of students. Problems with misclassification of students arose because of difficulties with the NCES definitions of student levels. Table 22 indicates that such misclassification occurs most frequently for first-year and second-year students.

Misclassification may occur for other reasons in addition to institutions having problems with the NCES definitions. About 23 percent of the institutions reported that some students were misclassified in the HEGIS report. The schools then estimated the magnitude of the misclassifications. Translating these estimates into overall impact on the accuracy of the population value resulted in the following. It should be noted that these data come from self-reports of the institutions and do not provide a true post-validation. Assuming both lack of awareness of some misclassifications and reluctance to acknowledge this misclassification, these data provide conservative estimates on the magnitude of the problem.

Type and Control of Institutions	Percentage of Students Misclassified
Public four-year	.3
Public two-year	1.5
Private four-year	.5
Private two-year	.5
Total	.7

Table 20

Percentage of Schools Reporting Estimation of Enrollments by Student Level

Student Level	Total Institutions (Who Estimate Enrollments)	Public		Private
		Four-year	Two-year	Four-year
All levels	46.9%	13.4%	50.0%	53.8%
Freshmen	6.2		12.5	
First-time freshmen	8.7		17.7	
Lower division students	4.5			11.8
Four-year students and beyond	9.7		19.8	
Part-time undergraduates	4.9	19.8		6.3
Graduate students	8.4	66.8		
Irrelevant response	10.8			28.1

(Taken from data reported in Peng, 1979)

Table 21

Percentage of Respondents Using Different Methods of Estimating School Enrollments

Basis of Estimate	Total Institutions (Who Estimate Enrollments)	Public Four-year	Public Two-year	Private Four-year
Departmental Reports	10.1	43.1	9.5	
Percent of Previous Reports	34.8	13.4	42.3	32.4
Simple Deduction	23.7	43.5	23.4	17.7
Degree Intent	4.5			11.8
Irrelevant Response	17.1		5.1	38.2
Don't Know	9.7		19.8	

(Taken from data reported in Peng, 1979)

Table 22

Percentage of Students Misclassified as Reported by Respondents

Student Level 1/	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Specific levels not specified	1.08%	.44%	1.47%		6.67%
First-year students	3.90	1.15	12.12	.02	
Second-year students	2.58	.46	8.59	.02	
Third-year students	.08	.46		.02	
Fourth-year students	.01	.46		.02	
Third- and fourth-year students	.04		.15		
Undergraduates in general	.02	.09			
Graduate	2/			2/	
Unclassified	.59	2.14	.79		

1/ Level as reported by respondents
2/ Specific percentage was not given by respondents

(Taken from Table 2-3 in Peng, 1979)

Multiple counts of students. About 23 percent of the institutions reported that some students had been counted more than once in their reports. The percentage of students counted twice ranged from .1 to 10 percent among institutions reporting double counts; the overall impact on the accuracy of the total enrollment figures was small. The results shown in Table 23 indicate that the total double counting was less than 1 percent. "However, when translated into actual numbers, the percentage may mean that several thousand students were counted more than once" (Peng, 1979, p. 2-12).

Omissions of students and inclusion of dropouts. Such omissions and inclusions can occur for a variety of reasons. A major reason is that institutions may fail or be unable to update their records for late registrations and dropouts. Some institutions permit students to register or to drop out until late in the fall term. If so, enrollment reports produced earlier in the term may not reflect the enrollment later in the year.

The 1977 Fall Enrollment survey that was the focus of Peng's post-validation study (1979) had a due date of 15 October 1977. This is an early date in the fall term for many institutions. Indeed, about 38 percent of the institutions indicated that some students were omitted from the report because of the late registrations. As can be seen in Table 24, such omissions occurred at all levels. However, the magnitude of these errors was small, less than one percent.

A problem related to the inclusion of late dropouts is the problem of omission of late registrants. About 44 percent of the institutions reported that some students who later became dropouts were included in the HEGIS report.

Recognizing that the early due date for the survey may pose problems for the institutions, the researchers asked whether the enrollment data would differ if the due date were set at 15 November instead of 15 October. About 33 percent of the institutions indicated that the enrollment would differ. Among these institutions, 80 percent reported the extra time would permit the inclusion of late registrants and the exclusion of dropouts, while the remaining 20 percent stated that it would permit more verification of their records.

In addition, the institutions were asked what would be a more convenient due date. Table 25 presents these results separately for the shortened card form and for the detailed form. The majority of institutions indicated that

Table 23

Percentage of Students Who Were Counted More than Once as Reported by Institution Personnel

Student Level 1/	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year 2/
Specific level not specified	.03		.12		
First-time freshmen	.83			1.80	
Part-time undergraduates	3/	3/			
Graduate students	.01	.04			
Unclassified students	.02			.05	

1/ Levels as reported by respondents
2/ No double-counts were reported
3/ Specific percentage point was unknown, but reported to be very small

(Taken from Table 2-4 in Peng, 1979)

Table 24

Percentage of Students Not Included in the Report Due to Late Registration, as Reported by Institution Personnel

Student Level 1/	Total Institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
Specific levels not specified	.19%	.15%	.41%	.09%	%
First-year students	.31	.05	.81	.14	
Second-year students	.10	.11	.06	.14	
Lower division	.15		.51		
Third-year students	.06			.14	
Fourth-year students and beyond	.07			.16	
Upper division	.11	.15	.29	.01	
Undergraduates	.002	.01			
Part-time students	.09		.19		
Graduate students	.46	.93		.29	
Unclassified	.26	.03	.81	.02	.07
First-professional	2/			3/	

1/ Levels as reported by respondents
2/ Percentage was not known
3/ The actual percentages were not reported by respondents

(Taken from Table 2-5 in Peng, 1979)

Table 25

Due Dates Acceptable for Fall Enrollment Report

Due dates	All institutions		Public				Private			
			Four-year		Two-year		Four-year		Two-year	
	Card	Detailed	Card	Detailed	Card	Detailed	Card	Detailed	Card	Detailed
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Oct. 15	75.4	58.9	78.3	71.0	67.4	56.5	80.8	53.4	78.1	76.0
Oct. 25	1.2	1.2			4.1	4.1				
Nov. 1	9.0	17.4	4.2	9.1	7.7	14.4	9.2	24.3	21.9	5.9
Nov. 15	6.0	9.4	12.2	12.2	7.3	5.6	4.0	9.1		18.1
Dec. 1	2.2	4.7			7.0	11.7	.5	2.7		
Dec. 15	.9	1.7	5.3	5.3		1.2		1.1		
Jan. 1	.5	.5					1.0	1.0		
Jan. 15	2.9	3.8		2.4			4.5	7.3		
May 1	.4	.4			1.4	1.4				
July 15		.5						1.1		
Unknown	1.5	1.5			5.1	5.1				
Sample Size	119		26		35		43		15	

(Taken from Table 2-16 in Peng, 1979)

~~79~~

15 October was acceptable for both forms. However, over 85 percent of the institutions reported a due date of 15 November as acceptable.

The due date for the 1978 Fall Enrollment survey was set at 15 November. The problems with reporting late registrations and dropouts have, presumably, disappeared for some but not all of the institutions. Another survey of the institutions would be needed to identify the magnitude of the problem. However, as shown in the next chapter, about 50 percent of the schools return their forms after 1 December.

Operational errors. Peng (1979) mentioned some administrative problems that affected the data such as failure to follow the instructions, failure to complete the forms, and computational errors. Table 26 displays the extent of such errors that were corrected during the editing operations for the 119 schools included in the survey. The editing and correcting of these forms is an important process that will be discussed in the next chapter.

The post-validation study also investigated whether the quality of the data was related to the data processing system. A comparison was made of the measures of net differences between the data published by NCES and by the "most accurate" data constructed in the post-validation effort. Results from these comparisons, as shown in Table 27, indicated that institutions with computerized systems provided data with a higher degree of accuracy than institutions with other systems.

The study did not, however, indicate the extent to which the involvement of the state coordinator affects the quality of the data by introducing or by eliminating errors. A determination of the role of the state coordinators could be undertaken through a study of the processing undertaken by these coordinators. Such a study would be most revealing if done in conjunction with the Fall Enrollment data collection effort.

Table 26

Frequency of Corrections on Each Line Item in the Fall Enrollment Report

Student level	N	F	Y percent
Full-Time Students			
A. Undergraduates, total	115	10	8.7
1. First-time freshmen	115	4	3.5
2. Other first-year	104	6	5.8
3. Second-year	112	5	4.5
4. Third-year	70	2	2.9
5. Fourth-year and beyond	67	2	3.0
B. Unclassified students, total	79	3	3.8
1. Undergraduate level	73	11	15.0
2. Post-baccalaureate level	32	7	22.0
C. First-Professional students	27	0	0
D. Graduate students, total	50	2	4.0
1. First year	50	5	10.0
2. Beyond first year	42	5	11.9
Total full-time students	119	7	5.9
Part-Time Students			
A. Undergraduates, total	104	7	6.7
1. First-time freshmen	94	3	3.2
2. Other first-year	88	5	5.7
3. Second-year	97	4	4.1
4. Third-year	59	2	3.4
5. Fourth-year and beyond	59	3	5.1
B. Unclassified students, total	85	6	7.1
1. Undergraduate level	83	11	13.3
2. Post-baccalaureate level	38	7	18.4
C. First-Professional students	17	0	0
D. Graduate students, total	48	1	2.0
1. First year	48	4	8.3
2. Beyond first year	38	5	13.2
Total part-time students	116	3	2.6
Total, All students	119	8	6.7

NOTE:
N = Number of schools with students in that level of classification.
F = Frequencies of changes or correction.
Y percent = F/N x 100
The percentages were unweighted.

(Taken from Table 2-10 in Peng, 1979)

Table 27

Percentage of Net Errors by Data-Compiling System for Selected Student Levels 1/

Student level	Data-compiling system	
	Computerized systems	Other systems
Total	.26%	- .85%
Full-time undergraduate, total	.06	- 1.63
Full-time first-time freshmen	.72	3.31
Other first-year full-time students	- .91	-11.36
Full-time unclassified students	1.63	-21.54
Total part-time students	.31	- 2.77
Sample size	61	58

1/ The percentage was computed as follows:

$$\sum_{i=1}^{J} w_i d_i \Big/ \sum_{i=1}^{J} w_i p_i \cdot 100, \text{ where}$$

w_i is the sample weight, d_i is the net difference between the HEGIS published data and the reconstructed validation data, and p_i is the published HEGIS data for the i^{th} institution in the group.

(Taken from Table 2-14 in Peng, 1979)

VI. Data Preparation

The editing of the HEGIS Fall Enrollment survey provides some quality control on the data collection operations. The procedures followed in the editing operation are described below. Later sections in this chapter provide some indication of the level of the quality of the data being submitted, the extent of editing and imputation, and the quality of the resulting data.

Description of the Coding, Editing, and Imputation Procedures

All of these operations are handled by a survey contractor. The steps in this process are depicted in Figure 4. As batches of forms are delivered to the survey contractor from NCES, they are grouped and logged. Then a scan edit prepares the forms to be machine-readable and identifies any glaring errors or inconsistencies. The manual edit attempts to identify and resolve such problems as missing totals. Although reports of such errors are prepared by the survey contractor for NCES, these reports are not retained. However, all the information on the results of the manual and machine edits are retained with the schools' HEGIS form at NCES or at OCR.

Following the manual edit, the forms are again logged and sent to keypunch. The keypunch operation achieves a level of accuracy of at least 99.5 percent. The listing of the keypunch output is checked for errors and then the forms enter the machine edit.

The machine editing follows the HEGIS Requirements and Specifications Manual, checking for internal consistency and for similarity with the previous year's data. The latter involves checks on presence or absence in certain data fields from one year to the next and checks on "tolerance levels." These tolerance levels for changes from one year to the next are set by the NCES staff. In general, however, decreases in enrollment have tighter tolerance intervals than increases in enrollment.

The final data files contain only the corrected data. Fields are left for the purpose of flagging imputed data. Data that were corrected at earlier stages and data that remain outside the tolerance levels do not receive any flags on the files. And, in reality, no data are ever flagged in HEGIS.

Editing steps. Prior to keying in a school's enrollment data, the survey contractor staff subject the raw data to a considerable degree of

Figure 4. HEGIS Processing by the Contractor

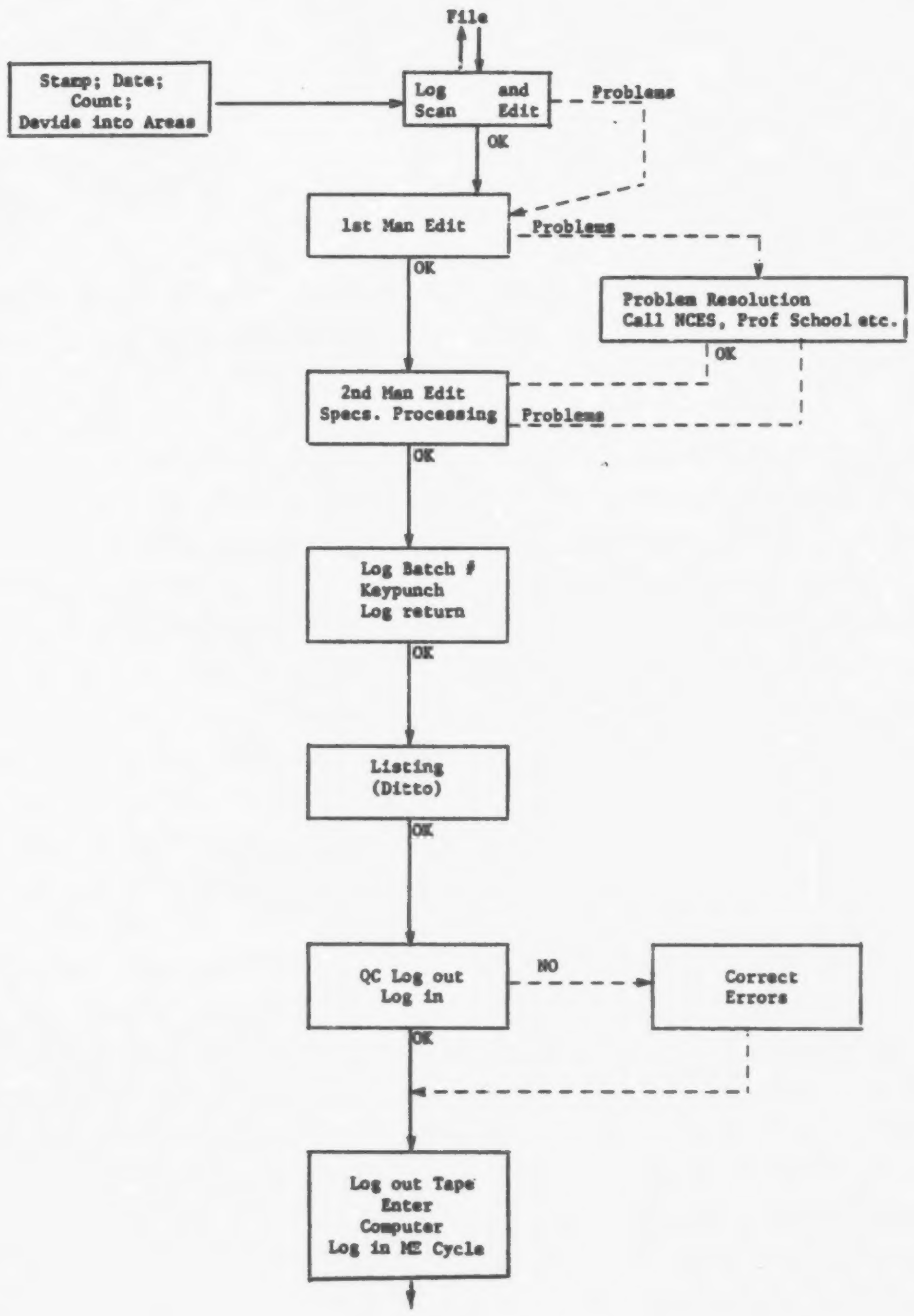
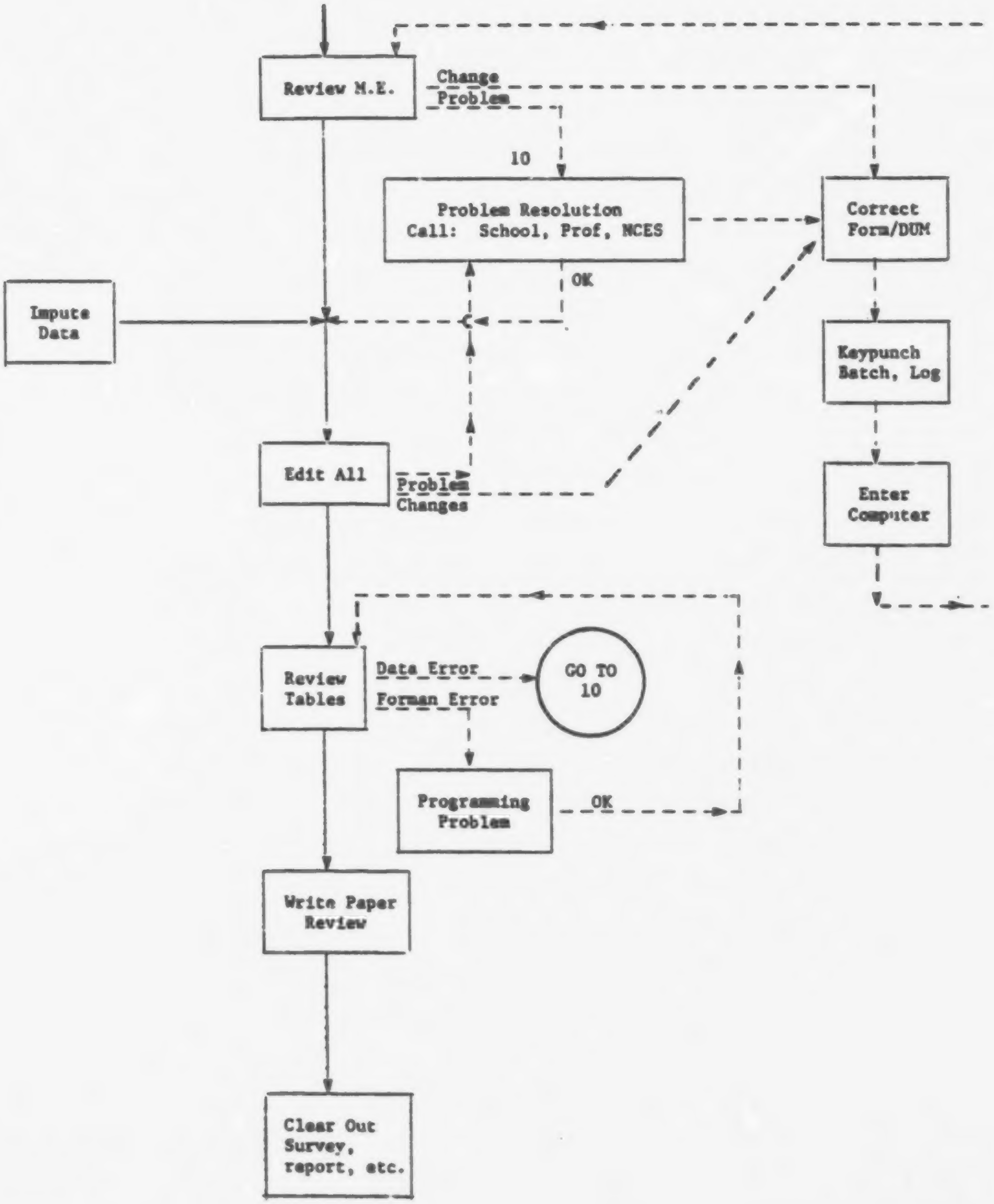

File
Stamp; Date; Count; Devide into Areas
Log and Scan Edit
Problems
OK
1st Man Edit
Problems
OK
Problem Resolution Call NCES, Prof School etc.
OK
2nd Man Edit Specs. Processing
Problems
OK
Log Batch # Keypunch Log return
OK
Listing (Ditto)
OK
QC Log out Log in
NO
Correct Errors
OK
Log out Tape Enter Computer Log in ME Cycle

Figure 4 (cont'd)


Review M.E.
Change
Problem
10
Problem Resolution
Call: School, Prof, NCES
Correct
Form/DUM
OK
Impute
Data
Keypunch
Batch, Log
Edit All
Problem
Changes
Enter
Computer
Review
Tables
Data Error
Forman Error
GO TO
10
Programming
Problem
OK
Write Paper
Review
Clear Out
Survey,
report, etc.

85

editing. Basically, the data are checked for internal consistency and for similarity with the previous year's data. Each school is subjected to the same procedure; that is, HEGIS does not single out a sample of schools for more thorough review.

The editing process centers around a so-called "machine edit." The machine (a computer program) does the consistency and similarity checks, and staff members then follow up on any discrepancies uncovered by the program. Typically, staff members call or write to the person supplying the data about any discrepancies which cannot be resolved by inspection of the form. This correspondence, as well as the results of the editing procedure, are documented on the "Edit cover sheet," which is attached to the raw data forms. If correspondence with the school is undertaken, this is documented in "phone logs" or "edit memos," which are also attached to the data forms. These three documents are permanently kept with the raw forms.

The consistency checks carried out by the program are of three types:

a) Add checks,

b) Presence/absence checks, and

c) Tolerance checks.

The add checks test whether enrollment figures actually add up. For example, the program checks that the number of undergraduates plus the number of graduates equals the total enrollment.

The presence/absence checks evaluate whether a school has reported data in categories that are inappropriate for that type of school (e.g., a two-year school that reports medical students), and whether different lines are filled in from one year to the next.

Finally, the tolerance checks test whether the total enrollment is within a tolerance interval as compared to the previous year's figure. These tolerance intervals are chosen arbitrarily by the director.

Of these three types of checks, the add checks are the most involved and definitive. An absolute requirement for entry onto the file is that the school's data eventually pass the add checks. However, a school can fail some of the other types of checks and still be included (e.g., if the total enrollment is found to be "out of tolerance" due to errors in reporting the previous years' data or due to a different way of defining enrollments, the unrevised figures may be entered onto the file.) But the enrollment figures that are reported must be internally consistent.

The add checks are applied to each page of the form, and the total page is tested against the sum of the detail pages. For each row in every page, the total for males and females separately must equal the sum of the racial/ ethnic breakdowns. For each column the following checks are done (see the Fall Enrollment Survey form in Figure 1, page __):

a) Rows 2,3,4,5,6 = row 1 [Undergraduates]

b) Rows 8,9 = row 7 [Unclassified]

c) Rows 12,13 = row 11 [Graduate students]

d) Rows 16,17,18,19,20 = row 15 [Part-time undergraduates]

e) Rows 22,23 = row 21 [part-time unclassified]

f) Rows 26,27= row 25 [part-time graduate students]

g) Rows 1,7,10,11= row 14 [full time students]

h) Rows 15,21,24,25 = row 28 [Part-time students]

i) Rows 14,28 = row 29 [grand total]

As mentioned previously, the final add check compares each entry on the total page against the sum of the corresponding entries on the detail pages.

If a school's data fail any of these add checks, the problem is resolved by a staff member. Certain types of simple, obvious errors are handled by the staff member. Other types of errors require contacting the school for clarification.

Procedures used to contact the schools. Whenever the staff cannot resolve the problems uncovered by the machine edit, they contact the school by telephone or in writing. It is not entirely clear how the staff decides between these two alternatives. If the school is called, the communication is documented on the "Phone Log" and summarized on the Edit Cover Sheet (Figure 5). If the communication is written, an appropriately filled out "Edit memo" is sent (Figure 6). This correspondence is also summarized on the Edit cover sheet.

The categories and alternatives in the Edit memo provide some insight into the way discrepancies are resolved. The first three categories pertain to problems with the add checks; categories 6, 8, and 9 pertain to tolerance checks; categories 4 and 10 refer to the presence/absence checks; and categories 5 and 7 deal with the problem of nonresponse.

As mentioned above, only problems that involve addition errors must be resolved. For presence/absence checks and tolerance checks, schools may (and do) respond with a comment or explanation rather than changing the

Figure 5

SAMPLE FECR EDIT COVER SHEET

HEGIS XIII EDIT COVER SHEET

FICE ____________

INST. ____________

PUB. SEQ. ____________

ACTION	DATE / INITL	FECR	Batch #
		PROBLEM DESCRIPTION	
PHONE RESPONSE			
MANUAL EDIT I			
PROBLEMS			
CALL TO INSTITUTION			
RESOLUTION OF PROBLEM			
EDIT MEMO SENT			
EDIT MEMO RETURNED			
RESOLUTION OF PROBLEM			
MACHINE EDIT		CYC NO.	MACHINE EDIT PROBLEM DESCRIPTION
MACHINE EDIT I			
EDIT ALL REVIEW			
CLEAN DATA BASE			

Figure 5 (cont'd)

HEGIS XIII PHONE LOG

Date Ext Used	Area Code	Phone Number	Person Contacted	Remarks - Resolution

Figure 6

EDIT MEMO FOR FECR SURVEY

____ Name of Institution

________________________ Institution FICE Code Number

Data in Question Because:

☐ 1. Line(s) __________ do not add to the total(s) in Column 13 and/or Column 14.

☐ 2. Column(s) __________ do not add to the total(s) on Line(s) __________, respectively.

☐ 3. The sum of the total students reported on the major field pages is higher/lower than the total figures on the summary page.

Line(s) ________________
Column(s) ________________

☐ 4. Students were/were not reported on Line(s) __________ of this year's form, but were/were not reported in those same categories in fall 1977. Please revise this year's report or explain the difference in the "Comments" section (item 13 below).

☐ 5. Students are reported on the total line(s) listed below but not broken down onto their detail lines. Please correct the designated items, with a brief explanation, if necessary, in the "Comments" section (item 13 below).

a. Total reported on Line __________ needs breakdown onto Lines __________.
b. Total reported on Line __________ needs breakdown onto Lines __________.
c. Total reported on Line __________ needs breakdown onto Lines __________.
d. Total reported on Line __________ needs breakdown onto Lines __________.

☐ 6. Your first-time freshmen enrollment for fall 1978 seems rather high/low as compared to fall 1977. Please correct or confirm with a brief explanation, if necessary, in the "Comments" section (item 13 below).

☐ 7. Unknown racial/ethnic data reported for 1978. Please correct using one or a combination of the following:

a. Distribute the unknown racial/ethnic data using some algorithm based on your institution's 1976 Fall Enrollment and Compliance Report response.

b. Prorate the unknown data based on the racial/ethnic distribution of the known data.

c. Any other method which would be appropriate to your institution.

Figure 6 (Continued)

☐ 8. The ratio of full-time equivalent (FTE) enrollment (Column 15) seems rather high/low as compared to the total part-time headcount (Part II) for Line(s) __________. Please correct or confirm the designated items with a brief explanation, if necessary, in the "Comments" section (item 13 below).

☐ 9. Your total enrollment for fall 1978 seems rather high/low as compared to fall 1977. Please explain briefly the reasons for the change in the "Comments" section (item 13 below).

☐ 10. Major Fields 1204, 1206, 1218, and 1400 should reflect only students at the first-professional level of enrollment. All other students should either be placed in Major Field 9000 (All Other) or be placed in another appropriate field.

☐ 11. Other Apparent Discrepancies: ______________________________

12. Editor's Comments: ______________________________

13. Comments by Respondent: ______________________________

(FOR NCES USE ONLY)

Memo prepared by ______________________________ Date ______________ figures themselves. However, the add checks and nonresponse problems must be resolved. As is made explicit in the alternatives to category 7, this sometimes involves imputing certain data points so that all appropriate cells are filled and that the filled-in data add up properly. Category 7 deals with nonresponse to the racial/ethnic categories within a line, while category 5 deals with nonresponse to detail lines within a column (e.g., unclassified total not broken down into undergraduate and graduate).

There are some instances when problems other than the add checks require that the originally submitted data be changed. These problems deal with data entered in enrollment categories inappropriate to that school. For example, a two-year college that reports medical school students would be required to move or otherwise modify these figures. This type of problem is generally flagged as category 10 or category 4. However, certain types of category 4 problems require only an explanation rather than data modification. For example, a two-year college reporting only freshmen one year and both freshmen and sophomores the next would be contacted because they reported students in a previously unused category. The school could simply respond that the data are correct.

Similarly, categories dealing with the tolerance checks may only elicit a comment or explanation from the school. The enrollment may change drastically from year to year because the method of computing the enrollment changes. If this is the case, the school is not required to use the same method as in previous years, or even correct previous years' figures if they are found to be in error.

In summary, the consistency checks used by HEGIS are much more stringent with respect to within-year criteria than the between-year criteria. Data which do not pass the add checks are, in general, simply not entered onto the final HEGIS tape; however, data which do not pass the tolerance and some of the presence/absence checks are entered onto the file after the school has been contacted and permitted to explain the problem.

In very rare instances, a school will simply not provide HEGIS with certain breakdowns, such as racial/ethnic data. If the staff is unable to secure numbers from the school in question, the data which has been obtained will be entered onto the file; blanks (zeros) will be inserted, denoting missing data, and the accompanying documentation will list each such school and describe the specific problem (Figure 7). This is done in preference to imputing the missing data.

Figure 7

HEGIS XIII
FECR MASTER FILE STATUS

The master file tape contains data for 3,170 institutions from a universe of 3,173 reporting units. Three institutions are not included because they have enrollments which are not applicable to the Fall Enrollment Survey for the following reasons:

1) San Francisco Community Colleges Center System (FICE 004502) has only non-credit adult education.

2) Center for Degree Studies, in Pennsylvania (FICE 004049) is comprised solely of correspondence students.

3) The Community College of the Air Force (FICE 012308) does not offer classes. Students taking classes through this program are included in enrollment counts for the institutions in which they are actually taking classes.

There are four institutions on the master file that contain arithmetic inconsistencies in order to maintain the June 1979 published totals.

1) Missouri Institute of Technology (FICE 002455). On the summary page, line 14, total full-time men (718 students) and total full-time women (22 students) are purposely omitted.

2) Mount Olive College, in North Carolina (FICE 002949). On the summary page, line 14, total full-time men (126 students) and total full-time women (199 students) are purposely omitted.

3) Polytechnic Institute of New York (FICE 002796). The total columns contain an additional 1,995 students (1,863 men and 132 women) not reflected in the detail columns. These additional students appear on the graduate students lines.

4) Baker Junior College, in Michigan, (FICE 004673). Line 16 contains 1,356 additional students (50 men and 1,306 women).

Six institutions provided total enrollments only, without breakdowns by specific year of study and racial/ethnic group.

1) Golden Gate Baptist Seminary, in California, (FICE 001204).

2) New Orleans Baptist Theological Seminary, (FICE 002019).

3) Washington Theological Union, in Maryland, (FICE 010065).

4) Mesivita Eastern Parkway Rabbinical Seminary, in New York, (FICE 009633).

5) Yeshivath Vizhitz, in New York, (FICE 013027).

6) Yeshivath Zichron Moshe, in New York, (FICE 011821).

Figure 7 (cont'd)

Those institutions which did not provide racial/ethnic data are:

1) Southern Baptist Theological Seminary, in Kentucky, (FICE 00 9R2).

2) Southeastern Baptist College, in Mississippi, (FICE 002435).

3) Saint Mary's Seminary-College, in Missouri, (FICE 002508).

4) Saint Michaels Passionist Monastery, in New Jersey, (FICE 002637).

5) University of the State of New York, External Degree Program, (FICE 011716).

6) Roanoke Bible College, in North Carolina, (FICE 029088).

7) Southeastern Baptist Theological Seminary, in North Carolina, (FICE 002963).

8) American College, in Pennsylvania, (FICE 029018).

9) Bob Jones University, in South Carolina, (FICE 003421).

10) Brooks Institute, in California, (FICE 001123).

The U.S. Service Schools are not required to report race data, nor are they included in the OCR tables. If they do report race data, it is put on file unedited.

Three institutions, University of Texas at El Paso (FICE 003661), Hampton Institute, in Virginia, (FICE 003714) and California College of Podiatric Medicine (FICE 001135), could not provide a breakdown by year of study for graduate students. Two institutions, Bayamon Central University, Puerto Rico, (FICE 010015) and Hebrew Union College New York Branch (FICE 004054), could not breakdown their unclassified students.

Imputation procedures. The processing specifications provide guidelines for imputation (Volume 2, pp. 1-59ff). The central guideline is that:

> data file must contain a 100 percent response by the institutions and 100 percent in terms of survey data cells appropriate for that institution. In most cases for the Fall Enrollment Survey, approval of the imputation must be sought from the institutions concerned (p. 1-59).

This is very obviously a stringent guideline in the sense that it places a great burden on the respondent. It, of course, assumes not only that every school in the universe is capable of providing the information, but also that these schools are all willing to do so. On the surface, it would seem unlikely that it is possible for HEGIS to achieve such cooperation from the entire universe of schools, especially in the years when the OCR data are collected.

As a result, a considerable amount of imputation is actually taking place. Regrettably, the amount of imputation being done is not preserved on the HEGIS file using the imputation fields. This is because the term imputation is officially defined so that the category is never used. According to the processing specs: "Imputation implies that no information was received formally or informally from the institution" (Vol. 2, p. 1-62). The definition goes on to distinguish between official imputation and "valid responses":

> Forms having data imputed by NCES personnel shall be considered valid HEGIS XIII responses in any of the following cases:
>
> (1) Imputation was conducted via telephone with the institution in question.
>
> (2) A copy of the imputed questionnaire was mailed to the institution, and concurrence was received by NCES.
>
> (3) A copy of the imputed questionnaire was mailed to the institution, revisions to the imputation were received by NCES and subsequently applied to the imputed form.
>
> (Vol. 2, p. 1-62)

In other words, the imputation is done either by the school or with the concurrence of the school. However, HEGIS refers to such imputed data as "valid data"; it is the policy of HEGIS never to do anything which would meet their definition of "imputation."

This is best illustrated by the fact that racial/ethnic categories are left blank in the few schools which do not cooperate with HEGIS. Contrary to the impression given in the processing specs, the fields are not imputed after staff have failed to obtain figures from these schools. In fact, it would be more realistic to eliminate the imputation fields from the HEGIS file because as a matter of policy they will never be used (Pepin, personal communication, 1981). Although the labelling of these data as "valid data" rather than "imputed data" does not alter the amount of error introduced, it does make it difficult to determine the prevalence of these procedures.

Potential Sources of Error

Errors in the keypunching operations. No data are available to indicate the level and type of keypunch error. The *Requirements and Specifications Manual* indicates that the error rate will be less than 5%. So we can assume that level as a maximum. With the further editing, it can be assumed that most, if not all, of these errors are discovered and corrected.

Errors in the editing operations. As a preliminary step, we redid the add checks on the HEGIS XIII file. Not surprisingly, these runs determined that all schools, save the few noted in the documentation, pass the add checks. (See Figure 7 for this listing.) Although this established that the editing procedures are "successful" in the sense that their goals are met, it does not establish that the data are error-free; clearly, the add checks constitute a necessary but not sufficient condition for error-free data. The "true" headcounts, presumably, have this property (e.g., assuming double majors are only counted as one person), but there are infinitely many other configurations which also have this property.

Errors Identified by Secondary Analyses

Further analysis was needed to identify the types of errors and the magnitude of these errors. The following sections on secondary analyses of the survey forms and edit materials and on the post-validation study provide these details.

To assess further the overall quality of the data turned in by the schools, and to evaluate the editing procedures used in HEGIS, we undertook a sample survey of HEGIS schools. Unlike the Peng post-validation study, which contacted the respondents themselves, this survey only examined archival data. We secured a sample (N=200) of raw HEGIS forms and accompanying edit materials from OCR. Along with the raw HEGIS forms, OCR retains the Edit cover sheet, the Phone log, and the Edit memo, if one has been sent. The HEGIS form itself shows the history of corrections to the data, and the accompanying edit materials describe the sequence of events in the editing process.

This survey was undertaken in order to address the following issues:

a) What is the level of quality of the data being submitted?

> That is, how frequently are the editing procedures actually needed?
>
> What proportion of schools turn in figures that do not pass the add checks?
>
> How often do schools list students in incorrect categories (e.g., third-year students in a two-year school) and how often do schools list students in programs they don't even offer (e.g., law school students listed in a four-year undergraduate school)?
>
> How well are the schools able to resolve students into the specialized categories on the HEGIS form (e.g., separating first-time freshmen from other first year and breaking down totals into the racial/ethnic categories)?
>
> Do the tolerance checks indicate that there are substantial fluctuations in total enrollments from year to year? If so, is it due to changing definitions or mistakes made in previous years?

b) How often are data imputed?

> In response to calls or letters from HEGIS staff, how frequently do schools impute data (as defined by the three criteria listed above)? Are some categories imputed more often than others?

c) Is the HEGIS system of editing capable of singling out those schools that do not use HEGIS definitions in computing enrollment figures?

> That is, is this system capable of finding these schools and getting them to use the HEGIS definitions?

The purpose of investigating the first issue is to determine the respondents' ability to do the task. Clearly, if a school is unable or unwilling to provide accurate data, no amount of editing will ever eliminate this as a source of error. The second issue, highly related to the first, is an attempt to roughly quantify the amount of imputed data present on the HEGIS file. Any estimate that we construct from this survey will obviously be conservative, because this procedure is only sensitive to imputations made as a result of being contacted by HEGIS staff; imputations made on the original form being submitted to HEGIS cannot be detected.

The third issue deals with the general problem that the HEGIS editing checks are for internal consistency rather than invalidity. That is, the HEGIS checks will fail if incorrect data are consistent; only if invalid data are also inconsistent will the checks be of value.

Procedures. To investigate this issue, we originally planned to secure the HEGIS forms from the schools sampled by Peng. As discussed earlier in this report, an appreciable percentage of these schools said that they did not use NCES definitions, had trouble reporting first time freshmen, estimated enrollment figures, and so forth. Our plan was to see whether these departures from the prescribed procedures were detected in the edits, and whether the edited data were closer to the reconstructed data provided in the post-validation study. Regrettably, we were unable to secure the list of these schools and the post-validation data tapes from the post-validation contractor. Thus, we were unable to address this issue.

As an alternate plan, we conducted analyses of a sample of available Fall Enrollment Survey forms and edit materials. The schools to be included in our survey were sampled from the HEGIS universe and chosen by the Lahiri method (1950; Cochran, 1977) described earlier in the section on the selection of institutions for the Early Release Survey. The raw forms were obtained from OCR and duplicated for the 200 institutions. These forms and the accompanying edit materials were then evaluated.

In an attempt to summarize the information in these materials, we devised the following categories of problems and scored each school in the sample according to presence/absence of these characteristics:

(1) Date returned. In view of the problems schools had with the October 15 deadline, we noted when the forms were actually submitted in 1978, the first year in which the deadline was extended to November 15.

(2) Math errors. If a school failed the add checks in either the manual or machine editing phase, HEGIS staff wrote "Math error" on the edit cover sheet and described the problem.

(3) Impute Page. Occasionally a school only turned in a total (9999) page. If so, HEGIS generated an "all other" page (9000) and set the values of this page equal to the entries on the 9999 page.

(4) Unknown students. If a school reported students whose race/ethnicity were unknown or a non-HEGIS category (e.g., "mixed" is reported in Hawaii), this was marked on the Edit cover sheet.

(5) Prorate. If, during the course of the editing procedure, the school or the HEGIS staff prorated students, this was noted in either the phone log or the edit memo along with the method used for proration (e.g., use last year's figures or distribute a total equally among the relevant categories).

(6) Inappropriate categories. If a school reported students in a category believed to be inappropriate to that school, this was noted on the cover sheet.

(7) Out of tolerance. If some figures were found to be out of tolerance in the machine edit, this was noted.

(8) Form not equal to tape. We checked whether the total line on the total page matched the data on the tape.

Results. The following tables report percentages of schools whose forms contained the above problems, broken down by public/private and two-year/four-year.

1. Date returned. Despite the more lenient deadline, fewer than half the schools were able to turn in their forms before 1 December. The following indicates the percentage of schools who submitted their forms later than the end of November.

	Submitted After 1 December
Public	
Four-year	58%
Two-year	44%
Private	
Four-year	48%
Two-year	60%

2. Math errors. Math errors were much more common among four-year schools, as the following tables make clear. This may have been due to the

fact that more numbers had to be reported and there were many more add checks that could be failed.

Table 28

Percentage of Schools Having Math Errors

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	22%	89%	54%	89%
Yes	78%	11%	46%	11%

3. Impute page. The converse of the situation with the math errors occurred with the impute page. A higher percentage of two-year schools had page 9000 (other) imputed. This happens when the school only turns in a total page, with no breakdowns.

Table 29

Percentage of Schools Having Page 9000 Imputed

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	97%	82%	94%	83%
Yes	3%	18%	6%	17%

4. Unknown students. There was a tendency for a higher percentage of four-year schools to report unknown students.

Table 30

Percentage of Schools Reporting Unknown Students

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	75%	96%	83%	94%
Yes	25%	4%	17%	6%

Because the policy is to report complete data, some adjustment must be made in these cases. All of the schools with unknown students were contacted by the survey-contractor staff, and we attempted to categorize the methods used to adjust the original figures. Two methods occurred frequently enough to emerge as distinct strategies. They are:

(1) Use some arbitrary method, such as distributing the unknown students equally across the HEGIS race/sex categories, or

(2) Use some data-based method, such as distributing according to the known proportions, or last year's report.

From some forms, it was not entirely clear what method had been used. The following tables give these breakdowns.

Table 31

Percentage of Schools Using an Arbitrary Method to Distribute Reported Unknown Students

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	87%	100%	67%	0%
Yes	13%	0%	33%	100%

Table 32

Percentage of Schools Using a Data-Based Method to Distribute Unknown Students

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	37%	50%	93%	100%
Yes	63%	50%	7%	0%

5. Prorate. Prorating is also done when a school is unable to break down total lines into their component parts. The following table gives the percentages of schools that prorated during the editing process. Like the set of tables on unknown students, these percentages represent very conservative estimates; they only estimate the amount of prorating done during editing and the amount of schools who REPORTED unknown students. Since schools know of the requirement for complete data, some unknown percentage impute or prorate some of the data originally submitted to HEGIS.

Table 33

Percentage of Schools That Prorated During Editing

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	75%	84%	88%	95%
Yes	25%	16%	12%	5%

As with the problem of unknown students, the same two methods emerged as the most common ways of dealing with the problem. A school that used an arbitrary method of prorating may, for example, agree to say that half of the unclassified students are undergraduates and half are graduates. HEGIS would then split the unclassfied totals in each column in half and enter these figures. Data-based methods include, for example, using the previous year's proportions.

Table 34

Percentage of Schools That Used an
Arbitrary Method to Prorate

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	87%	67%	50%	0%
Yes	13%	33%	50%	100%

Table 35

Percentage of Schools That Used a Data-
Based Method to Prorate

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	25%	67%	50%	100%
Yes	75%	33%	50%	0%

6. Inappropriate categories. Entering figures in inappropriate categories was slightly more common in two-year rather than four-year schools,

possibly because more categories are inappropriate. A common error was to enter third-and-fourth year students in a two-year school. Also, both four- and two-year undergraduate schools reported students in medical school, law school, dentistry, and veterinary medicine. This was rather puzzling, since these schools did not have any such programs.

Table 36

Percentage of Schools Reporting Students in Inappropriate Categories

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	87%	81%	96%	100%
Yes	13%	19%	4%	0%

7. Out of tolerance. A higher proportion of four-year schools were found to be out of tolerance, but the reason for this is not clear to us. This statistic is a bit hard to interpret, because the tolerance intervals are themselves arbitrary and are a function of school size. Perhaps they are too narrow for larger schools.

Table 37

Percentage of Schools Found to be Out of Tolerance

	Public		Private	
	Four-year	Two-year	Four-year	Two-year
No	41%	88%	69%	83%
Yes	59%	12%	31%	17%

There were no instances in which figures were revised on the basis of this check, despite the fact that all schools were contacted and responded to the inquiry. Schools explained that the difference was due to any number of causes, such as

(1) Enrollment changed due to changing financial situation and/or changes in college aged population.

(2) Method of counting had changed.

(3) More (or fewer) students are counted due to a different policy.

(4) Previous year's figures were wrong.

Although this edit check does seem to be uncovering real sources of error, it is somewhat distressing that no action had been taken in response to the check.

8. Form not equal to tape. In every instance we found that the final, revised HEGIS figures did, in fact, appear on the HEGIS tape. Thus, it is clear that these figures are being transcribed correctly.

Discussion. From this survey we learned several things about the amount and type of editing done on HEGIS forms as well as gaining an appreciation about the general quality of data being submitted to HEGIS. Several features of the editing process emerged.

First, it is very obvious that math errors are very common, especially among 4 year schools. Schools seemed to have special difficulty passing the final add check, namely that the entries on the total page are the sums of the entries on the detail pages. The usual style of the editors was to use the 9000 page (all other) to absorb the discrepancy. Thus, an assumption used in dealing with math errors was that the total page and the detail pages referring to specific majors are correct. This, of course, is only one possibility; any combination of the pages could contain errors.

This practice and assumption also was used in dealing with students entered in inappropriate major field categories. Frequently, the action taken is to move these students onto the 9000 page, leaving the remaining detail and total pages intact. And, by definition, the practice of imputing page 9000 when a school only turned in a total page makes the same assumption. Thus, the 9000 page appears to be functioning as a "dumping ground" for problems associated with inconsistencies and among the detail, all other, and total pages. Although the percentage of schools whose total enrollment was found to be out of tolerance is not of great meaning because the tolerance intervals themselves are arbitrary, the responses from the schools (contained on the Edit Memo) are meaningful. Their explanations of year-to-year fluctuations do shed some light on the problem of changing definitions of enrollment categories. HEGIS definitions are such that almost any method a school chooses to use is accepted. According to the

Edit memo, this includes any method used previously by the school, as well as "any other" method appropriate to the school. Some schools that were contacted responded by stating some real problem, such as use of a different definition or procedure (e.g., computer vs. manual) or even that the previous years' figures were wrong. Because the HEGIS editing system has not had a mechanism for correcting year-to-year inconsistencies, these sources of error are present in the final figures.

Finally, the results of this survey make it very clear that, in many cases, the HEGIS forms ask schools for more detailed information than they are capable of providing with available resources. The detailed breakdowns required in the OCR years, as well as certain breakdowns on each major field page, such as first-time freshmen and other first year, were especially difficult. Within a page, a school that could not resolve a total figure into its component parts sometimes left the lines blank. This would cause a failure of the add check and subsequent contacting of the school. At that point, a school was required to supply or impute these lines. From the documentation in the edit materials, it appeared that a frequent scenario was that a staff member might suggest a certain imputation strategy (e.g., half of freshmen are first-time freshmen) and see if a school would agree with that formula or suggest another. For inconsistencies across pages, it appeared that the most frequent suggestion was to ask if the "extra" students belonged on the page for the category 9000.

In other words, the fact that HEGIS meets its goal of 100% response rate from all schools in the universe and 100% complete forms from these schools does not accurately represent the overall quality of these data. In reality, a substantial amount of uncontrolled and undocumented imputation is done, and it appears that a substantial proportion of schools are at least partially unable to supply such detailed enrollment figures. This survey suggests that schools have an especially hard time resolving freshmen into first time and other first year and resolving unclassifieds into graduates and undergraduates. Also, page 9000 appears to be a dumping ground for assorted problems with the major field categories. We are somewhat puzzled as to why HEGIS has adopted the policy of not using imputation flags as a way of expressing to the user the relative degree of quality of the data. It would be more useful to secondary analysts if the imputation fields were used to denote any datapoints imputed by the school during editing. As with the post-validation study, we believe that grand totals are apt to be more accurate than individual breakdowns, and that some categories are more subject to error than others.

Errors Identified by the Post-Validation Study

The Peng (1979) post-validation study of the 1977-78 Fall Enrollment survey provides another source of data on the accuracy of the HEGIS data collection and editing operations. A sample of 120 institutions was selected for participation in the study. Only four institutions refused or were unable to cooperate. Substitutions were made for three of the four; the fourth institution, selected because of its continued refusal to participate in the Fall Enrollment survey, did not receive a substitution. Onsite interviews were conducted with institution personnel responsible for or involved in the preparation of the HEGIS report. Using information from internal and State reports, summary sheets, worksheets, and other written documents, data on enrollment as of 15 October 1977 were reconstructed. These data were compared with the published data. Discrepancies were resolved through follow-up telephone calls. It should be noted that about 15 percent of the institutions failed to provide documents other than the original HEGIS reports. These were assumed to be accurate since no other verification was possible.

The differences between the reconstructed and the NCES published data were computed for each of the 29 line items on the form (e.g., first-time freshmen). These differences were displayed in two forms:

- net differences, obtained as the sum of the actual differences between the reconstructed and the NCES data, and
- gross differences, obtained as the sum of the absolute value of the differences.

Table 38 provides the net differences between the reconstructed fall enrollment data and the published NCES data for each line item. Positive numbers indicate that the published data were over-reported, and negative numbers indicate that they were under-reported. For example, among full-time students, first-time freshmen were over-reported by 26,598, while other first-year students were under-reported by 32,696. Table 39 shows the net differences as a percentage of the published data. The results reveal that the greatest differences (or errors) occurred for other first-year students

Table 38

Net Differences Between NCES' Published Data and the "Most Accurate" Data 1/

Student level 2/	Total institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
I. Full-time students					
01 Undergraduates, total	-29,566	- 201	13,089	-44,818	2,373
02 First-time freshmen	26,598	1,063	27,896	- 1,911	- 451
03 Other first-year	-32,696	- 782	-31,710	237	- 441
04 Second-year	4,256	446	- 1,007	1,303	3,514
05 Third-year	- 2,332	- 1,397	--	- 935	--
06 Fourth-year and beyond	- 466	- 701	--	235	--
07 Unclassified students, total	- 4,716	- 87	- 6,198	1,813	- 245
08 Undergraduate level	- 3,206	60	- 5,505	2,484	- 245
09 Postbaccalaureate level	- 599	- 147	--	- 452	--
10 First-professional students	- 729	235	--	- 964	--
11 Graduate students, Total	2,302	1,632	--	671	--
12 First-year	733	39	--	694	--
13 Beyond first year	18	18	--	0	--
14 Total full-time students	10,939	1,573	8,742	468	156
II. Part-time students					
15 Undergraduates, total	33,412 3/	- 369	55,989 3/	-22,039	- 167
16 First-time freshmen	-16,436	760	-16,095	- 941	- 160
17 Other first-year	-39,813	84	-30,284	- 9,774	160
18 Second-year	- 3,698	326	- 3,554	- 304	- 167
19 Third-year	- 1,685	- 294	--	- 1,390	--
20 Fourth year and beyond	- 8,667	- 130	--	- 8,537	--
21 Unclassified students, total	-50,073	19	-72,875	22,582	200
22 Undergraduate level	-46,375	56	-70,144	23,513	200
23 Postbaccalaureate	2,637	- 253	--	2,890	--
24 First professional students	505	0	--	505	--
25 Graduate students, total	- 3,837	- 2,822	--	- 1,015	--
26 First-year	- 781	155	--	- 936	--
27 Beyond first year	- 1,471	- 1,393	--	- 79	--
28 Total part-time students	-18,763	- 1,980	-16,849	33	33
29 Grand total, all students	- 7,824	- 406	- 8,107	501	189

1/ Net differences are defined as the sum of the actual differences between the published and the reconstructed data. The sign "-" indicates that the published figure was smaller than the reconstructed one

2/ Details by type of institutions may not exactly add up to total due to rounding errors

3/ Some institutions reported only totals in the HEGIS report, and the totals were overestimated

(Taken from Table 2-6 in Peng, 1979)

Table 39

Net Differences as a Percent of the Published Data 1/

Student level	Total institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
I. Full-time students					
01 Undergraduates, total	- .49%	- .01%	.89%	- 3.48%	1.89%
02 First-time freshmen	1.61	.16	4.89	- .55	- .74
03 Other first-year	- 3.72	- .26	- 6.75	.27	- 2.79
04 Second-year	.28	.06	- .25	.39	7.40
05 Third-year	- .22	- .18	--	- .32	--
06 Fourth-year and beyond	- .05	- .10	--	.09	--
07 Unclassified students, total	- 2.30	- .12	- 6.14	5.64	- 816.67
08 Undergraduate level	- 1.97	.17	- 5.46	9.48	- 816.67
09 Postbaccalaureate level	- 1.40	- .40	--	- 7.63	--
10 First-professional students	- .32	.29	--	- .64	--
11 Graduate students, Total	.44	.43	--	.36	--
12 First-year	.27	.02	--	.97	--
13 Beyond first year	.01	.01	--	0	--
14 Total full-time students	.16	.04	.56	.03	.12
II. Part-time students					
15 Undergraduates, total	1.29	- .06	3.26	-10.70	- .67
16 First-time freshmen	- 2.66	1.28	- 3.06	- 4.06	- 1.77
17 Other first-year	- 4.67	.07	- 4.53	-20.97	2.23
18 Second-year	- .63	.26	- .85	- .79	- 1.99
19 Third-year	- .91	- .22	--	- 2.92	--
20 Fourth year and beyond	- 3.70	- .07	--	-17.11	--
21 Unclassified students, total	- 4.90	.01	-13.93	12.34	12.91
22 Undergraduate level	- 5.41	.03	-13.41	17.37	12.91
23 Postbaccalaureate	1.59	- .21	--	6.06	--
24 First professional students	2.71	0	--	2.88	--
25 Graduate students, total	- .64	- .63	--	- .65	--
26 First-year	- .22	.05	--	- 1.32	--
27 Beyond first year	- .59	- .85	--	- .09	--
28 Total part-time students	- .44	- .14	- .75	.01	.13
29 Grand total, all students	- .07	- .01	- .21	.02	.12

1/ The percentage was computed as follows:

$\sum_{i=1}^{J} w_i x_i \;/\; \sum_{i=1}^{J} w_i p_i \cdot 100$, where w_i is the weight, x_i is the actual difference, p is the published enrollment figure for the i^{th} institution, and J the number of institutions in the group.

(Taken from Table 2-7 in Peng, 1979)

and for unclassified students. Furthermore, although the percentage of errors for the grand total was small, the percentage of error for the more detailed categories was much higher. "Thus, most of the errors seemed to be caused by classification problems rather than uncertainty about the total enrollment" (Peng, 1979, p. 2-19).

Table 40 displays the gross differences between the two data sets, showing the total magnitude of error (both over- and under-estimation). In Table 41 can be found the gross differences as a percent of the published data. As with the data on the net differences, this shows that the greatest errors occurred with first-time freshmen, other first-year students, and unclassified students.

Table 40

Gross Differences Between NCES' Published Data and the Reconstructed Data 1/

Student level	Total institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
I. Full-time students					
01 Undergraduates, total	72,200	299	22,749	46,779	2,373
02 First-time freshmen	34,802	1,063	29,970	2,409	1,359
03 Other first-year	40,516	909	35,649	2,209	1,750
04 Second-year	9,608	507	4,210	1,377	3,514
05 Third-year	2,423	1,397		1,026	--
06 Fourth-year and beyond	2,466	701	--	1,765	--
07 Unclassified students, total	10,769	146	6,770	3,609	245
08 Undergraduate level	10,979	60	7,463	3,212	245
09 Postbaccalaureate level	796	179	--	616	--
10 First-professional students	3,089	235	--	2,853	--
11 Graduate students, Total	2,844	1,842	--	1,002	--
12 First-year	1,065	39	--	1,025	--
13 Beyond first year	308	308		0	--
14 Total full-time students	19,144	1,705	14,688	2,596	156
II. Part-time students					
15 Undergraduates, total	88,347	2,102	63,047	23,030	167
16 First-time freshmen	27,938	760	25,827	1,190	160
17 Other first-year	48,813	183	34,757	13,713	160
18 Second-year	11,189	388	9,173	1,461	167
19 Third-year	2,383	736	--	1,647	--
20 Fourth year and beyond	10,354	293	--	10,060	--
21 Unclassified students, total	127,552	3,507	101,081	22,764	200
22 Undergraduate level	128,376	670	103,811	23,695	200
23 Postbaccalaureate	6,082	3,193	--	2,890	--
24 First professional students	1,069	0	--	1,069	--
25 Graduate students, total	7,614	6,268	--	1,346	--
26 First-year	1,423	155	--	1,268	--
27 Beyond first year	4,917	4,838	--	79	--
28 Total part-time students	65,710	7,238	57,050	1,388	33
29 Grand total, all students	57,252	5,795	48,971	2,297	189

1/ Gross differences are defined as the sum of absolute differences between the published and the reconstructed data.

(Taken from Table 2-8 in Peng, 1979)

Table 41

Gross Differences as Percentage of the Published Data 1/

Student level	Total institutions	Public		Private	
		Four-year	Two-year	Four-year	Two-year
I. Full-time students					
01 Undergraduates, total	1.19%	.01%	1.55%	3.64%	1.89%
02 First-time freshmen	2.11	.16	5.26	.70	2.23
03 Other first-year	4.61	.30	7.59	.66	11.07
04 Second-year	.63	.07	1.04	.35	7.40
05 Third-year	.23	.18	--	.65	--
06 Fourth-year and beyond	.26	.10	--	.38	--
07 Unclassified students, total	5.24	.20	6.71	11.23	816.67
08 Undergraduate level	6.76	.17	7.40	12.26	816.67
09 Postbaccalaureate level	1.85	.48	--	10.40	--
10 First-professional students	1.35	.29	--	1.93	--
11 Graduate students, Total	.55	.55	--	.54	--
12 First-year	.40	.02	--	1.43	--
13 Beyond first year	.12	.22	--	0	--
14 Total full-time students	.27	.05	.94	.15	.12
II. Part-time students					
15 Undergraduates, total	3.42	.33	3.67	11.18	.67
16 First-time freshmen	4.52	1.28	4.91	5.13	1.77
17 Other first-year	5.73	.14	5.20	29.43	2.23
18 Second-year	1.90	.31	2.20	3.79	1.99
19 Third-year	1.29	.54	--	3.45	--
20 Fourth year and beyond	4.26	.16	--	20.16	--
21 Unclassified students, total	12.47	1.11	19.32	12.34	12.91
22 Undergraduate level	14.97	.34	19.84	17.50	12.91
23 Postbaccalaureate	3.68	2.71	--	6.06	--
24 First professional students	5.74	0	--	6.10	--
25 Graduate students, total	1.26	1.41	--	.87	--
26 First-year	.40	.05	--	1.78	--
27 Beyond first year	1.98	2.95	--	.09	--
28 Total part-time students	1.55	.52	2.55	.25	.13
29 Grand total, all students	.51	.11	1.28	.10	.12

1/ The percentage was computed as follows:

$\sum_{i=1}^{J} w_i d_i \,/\, \sum_{i=1}^{J} w_i p_i \cdot 100$, where w_i is the weight, d_i is the absolute value of the difference, p is the published enrollment figure for the i^{th} institution, and J, the number of institutions in the group.

(Taken from Table 2-9 in Peng, 1979)

VII. Data Analysis and Interpretation

Description of the Data Presentations

Data from the Fall Enrollment survey are presented in several different reports. The basic data and summary tables appear in the annual report entitled Fall enrollment in higher education. A brief section at the beginning highlights the enrollment statistics, showing total and percentage changes as compared with previous years. Following the brief discussion, summary tables are presented. These include data on individual schools. None of these tables display error rates. The appendices provide a one-page description of the data collection and editing and a copy of the questionnaire.

In addition to the above report, NCES prepares a brief paper on the results of the early release survey. This paper is titled Early release: Fall enrollment in colleges and universities: Preliminary estimates. The Early Release reports began to be published in Fall 1980. The report consists of a one-page summary followed by several pages of tables. The methodology section at the end of the report describes the sampling method and the potential sources of error, including a table of the estimated coefficients of variation.

Finally, OCR prepares a report based on analyses of the Fall Enrollment data "designed to assist OCR's regional and headquarters staff in their selection (or targeting) of institutions of higher education for compliance reviews" (Office for Civil Rights, 1979). The two major issues addressed by these analyses were:

(1) excess minority attrition in four-year undergraduate institutions (calculated by comparing the 1976-77 freshman minority enrollment in an institution with the 1978-79 junior enrollment within the same institution and by comparing the 1976-77 sophomore enrollment with the 1978-79 senior enrollment)

(2) underrepresentation of minorities and females enrolled in professional schools (i.e., medicine, dentistry, law, and veterinary medicine).

The report presents a ranking of the institutions for each component of the two compliance issues. The report provides no description of the data collection procedures and no indication of possible errors.

Potential Sources of Error

None of the reports discuss nonsampling error. This leaves the reader with the impression that the data are error-free. As has been shown in this present report, the data are not error-free. Some discussion of these errors should appear in the reports.

The report on *Fall enrollment in higher education* presents data on differences involving changes over time. Unfortunately, no checks or tests are performed. Thus, some of these statements discuss changes which are probably not statistically significant.

The *Early Release* report identifies two sources of error. First, enrollments were estimated for new institutions that had been added since the previous year. Previous years' average enrollments were used to estimate the total enrollment for these institutions. In addition, sampling error was discussed followed by a presentation of the sampling error as the variation among estimates from all possible samples. The report then states "we are approximately 95 percent confident that the actual enrollment figures which would be obtained from all 3,190 institutions are within two standard errors of the figures shown in the Summary Tables" (National Center for Education Statistics, 1980).

The OCR *Targeting Analyses* report does present the data on excess minority attrition based on tests of statistical significance. "Only institutions that were detected as having a statistically significant different racial enrollment distribution (with a 95 percent level of confidence) were retained and included in the final set of targeting reports..." (Office for Civil Rights, 1979, p. B-2). These were not described, however. The data on enrollment in professional schools was basically enumerative "since no statistical test of significance was performed on these data." In neither case is there evidence to indicate that differences in the rankings of institutions were significant.

VIII. Conclusions and Recommendations

Overview

The present report provided a process or error profile of the Fall Enrollment Survey within the Higher Education General Information Survey (HEGIS). It documents sources of error in the survey and identifies potential sources of error. In addition, it can serve as a model for future efforts. The following paragraphs summarize the findings and recommend future efforts.

Sources of Error

Major steps in the survey were identified. Information was then assembled to describe the processes undertaken in each step and to document the sources and magnitude of errors associated with each step.

Step 1: Survey objectives. The HEGIS, as an entire package, appears to be meeting the information needs of numerous federal and state agencies, associations of higher education institutions, and research organizations. Furthermore, a process exists for modifying and updating the survey objectives to meet information needs more closely.

Quantifying the "error" associated arising from a mismatch of information needs and survey objectives is impossible without detailed documentation on the matches and mismatches. We can identify some problem areas where such mismatches can be reduced or eliminated.

- Timeliness of HEGIS data is seen as a major problem. The Early Release Study addresses this problem, but the issue of data availability remains.

- More data are wanted on student characteristics and financial aid. As shown in our analysis of the data analysis and processing operations, many institutions are unable or unwilling to provide the data currently requested. Additional data demands on all the institutions would increase the burden and possibly reduce the accuracy. If such data are needed, they should be gathered in special samples of the HEGIS universe.

- Institutions are concerned about using the HEGIS data for comparison purposes. This was discussed primarily in terms of the financial data. It is, however, applicable to the enrollment data--given the lack of timeliness in the responses, the lack of similar definitions or interpretations of the given definitions, and the lack of appropriate detail. The major use of institutional comparisons with the Fall Enrollment data appears in the OCR analyses.

Step 2: Sampling. The HEGIS Fall Enrollment Survey involves the universe of institutions of higher education within the United States and its outlying areas. The errors in the frame are small, less than one percent of the universe. The errors result from undercoverage of certain institutions (i.e.. time lag between receipt of accreditation and entrance into the frame, refusal to cooperate), from overcoverage of certain institutions (i.e., time lag between loss of accreditation and removal from the frame, inclusion of vocational postsecondary institutions), and from misclassification. The effect of these frame errors on the enrollment figures is extremely small, probably less than .1 percent.

Step 3: Measurement instruments. The Fall Enrollment Survey uses alternate forms for alternate years of the survey. During the years when data are collected for OCR, categories for race/ethnicity and type of course are added to those of degree credit versus non-degree credit, resident versus extension student, grade level, full-time versus part-time student, and sex. The following list indicates the sources of error that were discussed.

- Uncertainty exists among the respondents as to the focus of the survey--the peak enrollment, the total number of students ever enrolled, the number of students currently enrolled, the students who completed the fall term, or the total number of students after the drop/add period. This uncertainty leads to inconsistencies in reporting among institutions. For example, the post-validation study found that about 56 percent of the institutions deleted dropouts from their reports, while the remaining 44 percent included them.

- Almost 25 percent of the institutions reported that they did not follow the HEGIS definitions for student levels, citing such problems as inability to distinguish first-time freshmen from other first year students, difficulty in classifying transfer students.

- Institutions tended to overreport full-time students and to underreport part-time students. The resulting error was less than one percent.

- Another problem involved that of interpreting and applying one of three definitions for the full-time equivalent (FTE) counts. No data are collected on the specific method used by each institution.

The chapter concluded with some suggestions for improvements in the wording and format of the questionnaire.

Step 4: Data collection. The data collection effort involves several different parties. NCES distributes the questionnaires to the institution; in most cases this is done through a state coordinator. Institutions compile the information through computerized systems (39.7 percent), through manual operations (30.9 percent), through aggregation of department reports (.4 percent) and through partly computerized and partly manual operations (16.0 percent). The responses are returned to NCES and then transmitted to the outside contractor for processing. Some potential sources of error include the response problems described in the section on measurement instruments, as well as the following:

- The post-validation study found that about 13 percent or the institutions had to estimate the enrollment counts. Information as to which institutions estimate enrollment counts and what methods are used is not gathered as part of the survey.
- Institutions acknowledged that the following errors occurred: misclassification (23 percent), multiple counts (23 percent), omission of late registrations (38 percent), and inclusion of late dropouts (44 percent)
- Institutions with computerized systems provided more accurate data than those with other systems, according to the results of the post-validation study.

The discussion below on the data preparation identifies further sources of error related to both the measurement instruments and to the data collection.

Step 5: Data preparation. The coding, editing, and imputation procedures are handled by an outside contractor following the operations outlined by NCES in the Requirements and Specifications Manual. Potential sources of error were identified.

- At least 20 percent of the institutions were unable to complete the forms without including math errors, imputing or prorating to indicate category breakdowns, or reporting students in inappropriate categories. In addition, a similar percentage of institutions failed the tolerance checks. None of these problems are recorded on the data tapes or in the accompanying documentation.
- Data from the post-validation study indicate a small percentage of error (less than one percent) for the total enrollment counts. Much higher level of error occur in the detailed categories.

These data suggest that a substantial percentage of the institutions are unable or unwilling to provide detailed category breakdowns. Because of the

HEGIS policy to have 100% response, these institutions are forced to estimate, prorate, and impute enrollment counts. Since none of this activity is recorded on the data tapes, one is left with the impression that no imputation occurred.

Step 6: Data analysis and interpretation. Three major reports are prepared using the Fall Enrollment data: (1) Fall enrollment in higher education, (2) Early release, and (3) OCR's Targeting analyses. None of these reports includes any discussion of nonsampling error, leaving the reader to assume that the data are error-free.

References

American Council on Education, Center for Research and Development in Higher Education at Berkeley, Education Commission of the States, National Center for Higher Education Management Systems at Western Inter-State Commission on Higher Education, & State Higher Education Executive Officers Association. Eighth Annual Conference on Higher Education General Information Survey (HEGIS). Washington, D.C.: American Council on Education, 1972.

Andrew, L. D. Analysis of uses of HEGIS data. Blacksburg, VA: Virginia Polytechnic Institute and State University, 1980.

Berg, P. Personal communication, 1982.

Brooks, C. A., & Bailar, B. A. An error profile: Employment as measured by the Current Population Survey (Office of Federal Statistical Policy and Standards, Statistical Policy Working Paper 3). Washington, D.C.: U.S. Department of Commerce, 1978.

Brown, K. G., Padgett, D. W., & Embry, L. R. Uses of national data systems by institutional researchers: Implications for the 1980's. Twentieth Annual Forum of the Association for Institutional Research, Atlanta, 1980.

Cochran, W. Survey Sampling (3rd ed.). New York: Wiley, 1977.

California State Department of Education. Directory of California private postsecondary educational institutions and public and private institutions "Approved for Veterans," 1978. Sacramento, CA.: Author, 1978.

Drews, T. H. HEGIS: A survey program becomes a system. Seventh annual conference on Higher Education General Information Survey (HEGIS). Washington, D.C.: National Center for Education Statistics, 1971.

Hyatt, J. A., & Dickmeyer, N. Joint study group on the utility of HEGIS finance data for institutional and higher education sector comparisons (Draft Report). Washington, D.C.: National Center for Education Statistics, American Council on Education, National Association of College and University Business Officers, 1980.

Lahiri, D. B. A method of sample selection providing unbiased ratio estimates. Bulletin of the International Statistical Institutes, 1951, 33, 133-140.

National Center for Education Statistics. Early release: Fall enrollment in colleges and universities, 1980: Preliminary estimates. Washington, D.C.: Author, 1980.

National Center for Education Statistics. Report of the eleventh annual conference on the Higher Education General Information Survey (HEGIS). Washington, D.C.: U.S. Department of Health, Education, and Welfare/ Education Division, 1975.

National Center for Education Statistics, Western Inter-state Commission on Higher Education, & American Council on Education. *Seventh annual conference on Higher Education General Information Survey (HEGIS)*. Washington, D.C.: National Center for Education Statistics, 1971.

Office of Civil Rights, & Value Engineering Company. *Targeting analyses in institutions of higher education based on excess minority attrition in undergraduate institutions and the underrepresentation of minorities and females in professional schools*. Washington, D.C.: Authors, 1979.

Peng, S. *HEGIS post-survey validation study* (2 vols.). Rockville, MD: Westat, 1979.

Pepin, A. J. *Fall enrollment in higher education 1978.* Washington, D.C.: National Center for Education Statistics, 1979.

Pepin, A. J. Personal communication, 1981.

Peterson, R. J., & Davis, G. C. *Education directory, colleges & universities 1979-1980*. Washington, D.C.: National Center for Education Statistics, 1980.

Podolsky, A. *HEGIS XIII 1978-1979 requirements and specifications manual* (2 vols.). Washington, D.C.: National Center for Education Statistics, 1979.

Roark, A. C. Washington's "most effective" lobbyist for higher education. *The Chronicle of Higher Education*, October 6, 1980, p. 3.

Russ-Eft, D. F. *Validity and reliability in survey research* (SAGE Technical Report No. 15). Palo Alto, CA: American Institutes for Research, 1980.

Topping, J. *Evaluation of the IEP costing procedures: A pilot study of six major research universities*. Boulder, CO: National Center for Higher Education Management Systems, 1979.

APPENDIX A:

Background Information on HEGIS

Needs for Information on Higher Education

Institutions of higher education are an important component within the American educational system. From the earliest reports on the condition of education, certain information about these institutions was identified as critical:

The names and locations of these institutions

The names of their presidents

The dates institutions were established

Their religious affiliations

The number of instructors on the staff

The number of students enrolled
Level of these students
Numbers of males and females

The tuition per month and the board per month at each institution

The number of volumes in the library

A detailed version of these information needs appears in the *Report on the eighth annual conference on Higher Education General Information Survey* (1972).

Over the years the information needs on institutions of higher education have increased. The following are some examples at the federal level. The National Science Foundation has used information on the enrollment counts for advanced degrees for planning support programs. The Accreditation and Institutional Eligibility Staff of the Bureau of Higher Education needs information on institutional characteristics to determine eligibility for participation in federal programs. The Department of Labor needs data on tuition for computing indexes of education cost and cost of living. They also need data on enrollments and degrees earned for preparing manpower reports.

State agencies, research organizations, and associations of higher education institutions also need data regarding higher education. These organizations need information on enrollments, faculties, finances, and staff to develop policies and recommendations. For example, the Mid-West Advisory Commission on Higher Education needed data on the residence and migration of college students to formulate a uniform policy for out-of-state students.

Purpose of HEGIS

The original purpose for HEGIS was to acquire and disseminate, to agencies within the Federal government and to the higher education community, more meaningful statistics on the operations of higher education (Drews, 1971). The meaningfulness of these statistics derived from the content as well as the format of the data.

Having more meaningful statistics was not the only purpose for HEGIS. As mentioned above, the response burden on institutions had been growing steadily. A secondary purpose for HEGIS was to reduce this burden. Related to the reduction in the response burden was the goal of attaining some consistency (across surveys and across years) in the data collection effort.

Method for Developing and Elaborating the Purpose of HEGIS

The HEGIS package was developed at the request of the higher education community (Drews, 1971). This community consists of the institutions of higher education, the various voluntary organizations within these institutions, professional groups within the institutions, and the federal, state, and local governmental agencies involved with or assisting higher education. Indeed, representatives of those various groups gathered at the Office of Education in the early part of 1966 to assist in developing HEGIS.

The higher education community continues to advise NCES in the planning for HEGIS. This is accomplished through the annual HEGIS planning conference. An excellent example of the recommendations made by the higher education community appear in the Final Report of the seventh annual conference on Higher Education General Information Survey (1971).

Match between the Information Needs and the Purpose of HEGIS

One potential source of error arises from the match or mismatch between the information needs and the purposes of the survey. There is a close match between the information needs and the purposes of HEGIS. The primary purpose of acquiring and disseminating meaningful statistics on higher education meets the needs of numerous federal and state agencies, associations of higher education institutions, and research organizations. Evidence for this statement comes from the report by Andrew (1980) titled Analysis of uses of HEGIS data. The report examined the questions of who uses the HEGIS data and for what purpose. It summarized the results of the following activities: (1) a conventional literature review to determine the trends in HEGIS usage, (2) a statistical sampling of the relevant literature to determine the level of use, and (3) interviews and surveys of users to obtain information on the use of the HEGIS data and on opinions about its use. The following conclusions resulted concerning the information needs served by HEGIS:

> HEGIS data have provided a foundation or base for the majority of reports and books that have affected public policy on higher education. Almost everyone that was interviewed agreed with this hypothesis while admitting to the principal investigator that it is difficult to show a direct cause and effect relationship. As noted more extensively in the body of the report, many factors and interests contribute to the development of public policy, not least of which is the lobbying of representatives of higher education. During the process of setting policy and making law, lobbyists and analysts both at the executive and legislative levels have to consider the interests of many constituencies and conflicting priorities. However, it appears from a review of higher education as well as from other literature that ideas behind much policy and law generally precede the full development of policy and its conversion into law by several years. In higher education, for example, the Carnegie Commission for the Study of Higher Education has produced extensive studies on higher education, many of which utilized statistics from the Higher Education General Information Survey system and other sources, such as the Census, to describe the condition of higher education and to provide a foundation for policy recommendations. It seemed to this author and to many interviewees that a considerable amount of higher education law and policies in the seventies appeared to be derivatives of much of what was recommended by such foundations as Carnegie and the research sponsored by various federal and state agencies. Other evidence that HEGIS data provide a base for law is found in the extensive quotation of HEGIS data during Congressional Hearings on Higher Education (see the review of literature for examples) and reports by interviewees. Most educational associations develop voluminous reports on the condition or projected condition of higher education for their own constituencies, as well as appropriate staff of congressional committees and executive agencies. In addition, the staff of associations and of Congress work closely together by telephone and memoranda with association

staff supplying data or analyses. The data come from the associations' own research, the Bureaus of Census and Labor Statistics, and from HEGIS. (For an example of how associations work with Congress, see Roark, Oct. 6, 1980, p. 3.)

Enrollment and financial data are used much more extensively than other survey data for analyzing the Condition of Higher Education, policy analysis, and for making decisions at state and local levels. This is probably true. (True is used here and elsewhere in the report in a relative sense.) However, Degrees and Other Awards Conferred data are used extensively in conjunction with enrollment data for manpower planning and evaluating affirmative action programs and persistence of students. Faculty and employee salary data is reported extensively as is tuition and fees because of the impact on personal and institutional decisions. These data are used to some degree in policy development.

Accuracy has improved. Generally the accuracy of all surveys is deemed acceptable. The lone exception to this is in aspects of the financial survey. The financial survey file is probably used more than other files in making complex analyses of the condition of higher education. Moreover, there are many difficulties in reporting and interpreting financial data because of differences among institutions in government and accounting practices. Thus, reports of dissatisfaction with the relative accuracy of the HEGIS file were not unexpected. The major problems with the financial file are summarized in Chapter II. The findings were drawn from Hyatt and Dickmeyer, An Analysis of the Utility of HEGIS Financial Data, May 22-23, 1980. It seems that many of the problems with the file would probably be corrected by more extensive documentation about the accounting practices and governance of certain institutions.

What was unexpected was the relatively high esteem that surveyees and interviewees had for the accuracy of most of the files. A recent study by NCES confirms the opinion of surveyees and interviewees about the relative accuracy of enrollment and degree data. The NCES study (Westat, 1979) reported that there was less than one percent difference between survey and audit data on enrollment and degree data. However, certain caveats are in order about the accuracy of the files. Some researchers are concerned about the levels of aggregation in the files on Enrollment and Degrees Awarded. Another respected researcher believes that the financial file is more accurate than perceived, relative to the other files, and that the concern about the file is a function of its extensive study and use, as she believes expectations concerning accuracy increase with the use of data. It is also worth noting that one interviewee familiar with how library data have been collected or estimated in the past questioned the accuracy of this file. Library and facilities data have not been reported nor collected for some time and, therefore, not used extensively, at least for complex analysis, in the last few years.

> HEGIS is a system that would have to be invented if it were not already in place because of the increasing need for data in policy making and planning. Everyone agreed with this notion.
>
> The uses of HEGIS data have increased significantly in recent years, particularly in the sophistication with which they are used.
>
> (Andrew, 1980, pp. iv-xv)

In addition to these conclusions, the Andrew (1980) report provided some tabulations that indicate the usage of the HEGIS data. Andrew's literature review identified four major uses of the HEGIS data: (1) the description or condition of higher education, (2) policy and planning decisions, (3) projections, and (4) other. Table A-1, taken from the Andrew (1980) report, indicates the frequency of use by different types of users. Description of higher education, including information on enrollment, accounts for the most frequent use of the HEGIS data (51% of the 873 citations). This use probably overlaps with that of policy and planning (22%), since most descriptions are produced for policy and planning purposes.

Further detail on the uses and purposes of HEGIS was provided by a survey of 109 data users. Table A-2 presents rank orderings of the purposes for the HEGIS use by types and users. Manpower planning and enrollment projections were ranked as first or second by all types of users. The importance of enrollment data is underscored by the results on the use of the HEGIS data. The highest percentage of respondents (65%) reported using the data from the Fall Enrollment Survey.

This report did, however, note some problems with HEGIS that relate to the major purpose of acquiring and disseminating meaningful statistics on higher education. These are listed below.

> Timeliness of HEGIS data is seen as a major problem. This was found to be a major problem with HEGIS. The delay of nearly a year or more, justified or not, between collection and distribution of data in machine processable form and hard copy publications is seriously affecting the use of HEGIS. Though there has been recent improvement in releasing tapes of certain files faster, there is still considerable dissatisfaction with the timing of releases. This dissatisfaction is reflected in findings from surveys and in the comments of researchers who work both for educational associations and institutions, charged with reporting to their constituencies and/or supplying data for

Table A-1

Purposes for Which HEGIS Data are Used by Groups Utilizing HEGIS Data*

(Numbers in parantheses are percentages [illegible] total)

	Purpose of Use				
User	Description	Policy/ Planning	Projections	Other/ None	Row Total
Federal Governmental Agency	16 (1.8)	7 (0.8)	2 (0.2)	7 (0.8)	32 (3.7)
State Government Agency	22 (2.5)	14 (1.6)	0 (0)	1 (0.1)	37 (4.2)
Quasi-Governmental Agencies	44 (5.0)	11 (1.3)	2 (0.2)	3 (0.3)	60 (6.9)
Institutions	205 (23.4)	108 (12.4)	26 (3.0)	98 (11.2)	437 (22.1)
General Public/ Other	106 (12.1)	38 (4.4)	11 (1.3)	38 (4.4)	193 (22.1)
Scholars	36 (6.4)	14 (1.6)	5 (0.6)	39 (4.5)	114 (13.1)
Column Totals	449 (51.4)	192 (22.0)	46 (5.3)	186 (21.2)	

Total = 873

*Most common purpose for using HEGIS data is for description; 22% of citations involve using HEGIS data for policy/planning.

(Taken from Table 2.1 in Andrew, 1980)

Table A-2

Uses of HEGIS Data Rank Ordered Within Groups by Purpose (1=most used)

	Rank Order of Frequency				
Purposes	Quasi-Governmental and Association	State Boards	Institution/ Scholars	Federal Government	Private Enterprise
(1) Investigating financial conditions of higher education	2	2	3	1	---
(2) Manpower planning	2	2	1	2	---
(3) Enrollment projections	1	1	2	2	1
(4) Market planning analysis	3	---	1	---	1
(5) Library planning	4	3	3	---	---
(6) Facilities planning	4	1	---	---	---
(7) Status of higher education					
(a) by private sector	1	1	3	2	---
(b) by public sector	2	1	3	2	---
(c) in adult and continuing education	3	3	---	---	---
(d) in vocational/ technical education	4	3	---	---	1
(8) Other	3	---	2	2	---

(Taken from Table 4.10 in Andrew, 1980)

making administrative and budget decisions. Students of higher education also voice the same complaint. The lack of timely data, as well as difficulties in accessing data in machine processable form (if the data aren't used regularly), probably leads institutions and associations to do more collecting of data through their own surveys (formally or informally) that would be unnecessary if HEGIS data were released more quickly.

However, the expectations of some institutional researchers for delivery of data to support budget proposals, etc., can probably not be met. The primary purpose of HEGIS was and is to report on the condition of higher education at the national level, though such reporting necessarily requires analyses of various sectors of the enterprise. But the data are also used for secondary purposes (for example, making comparisons among institutions by institutions and state agencies). These uses have occurred because the system provides for consistency in reporting on such matters as finances, degrees and enrollment for a universe of institutions. Generally, comparative data are wanted by state agencies and institutions for budget analyses. Since the budget cycle is almost continuous at the institutional level and budget development for the next year generally begins before actual data on the current year are collected by HEGIS, institutions find that they are required to use projections and revise them as actual data is collected. These revisions quite often are occurring as their reports to HEGIS go forward to intervening agencies, such as state boards, for edits and eventual forwarding to NCES for further edits. Thus, by the time NCES has the data for edit, institutions may have completed their budgeting process for the next year. The cycle and the process therefore appear to preclude NCES' ever delivering reports in time to support budget requests by institutions. Thus, what is going on will probably continue, and, in a sense, provides a use of HEGIS in a very informal way--the trading back and forth of data among institutions that they have collected for their own management or for HEGIS long before such data do, or could possibly, appear in HEGIS reports.

This is not to to excuse HEGIS from the requirement to report results of its surveys earlier. Currently, certain HEGIS data are reported in hard copy form as much as two years after the data were collected. Tapes and publications tend to be released as much as a year or longer after the data were collected. This is unacceptable. There was general consensus among interviewees that the data should be published both in machine processable and hard copy from six months and a year (even if this meant leaving out late reporting institutions, thereby sacrificing completeness and accuracy) after collection.

HEGIS data have not been used as extensively as they might be in reporting on the condition of women and minorities in higher education because overhead or start-up costs in using HEGIS data for analysis is relatively high. Experienced users tend to disagree that start-up costs are high; but then they have already

> paid those costs. There has been a spurt of studies on ethnic groups and women in higher education in the last year, quite a bit of it being published and disseminated since the review of the literature was published. Thus the conclusion may not be tenable in the future.
>
> *HEGIS is not being used as fully as it might be for policy analysis, planning and evaluation by either business or university scholars*. As noted earlier, there is only a small coterie of scholars and students in universities that is using HEGIS for the above purposes. While there are strong indications that data are being used somewhat by businesses for planning recruitment and evaluating or negotiating affirmative action programs, these uses seem fairly unsophisticated. There is little information in the general literature on higher education about the contents of HEGIS and how to use it.
>
> *More data are wanted on student characteristics and financial aid*. Without question more information is wanted on the latter. There appears to be more disapproval than approval for HEGIS' collecting data on student characteristics, institutional quality and outputs. However, there is more and more demand for such data from policy makers and consumers. Data are being gathered and data bases are in place or being developed. Some interviewees suggested that NCES should act as a broker in gathering data from other Department of Education program offices, funding the collection and maintenance of data bases, and disseminating data.
>
> (Andrew, 1980, pp. iv-xv)

The Andrew (1980) study also indicated the partial achievement of the secondary purposes of HEGIS--to reduce the response burden and to attain consistency.

> *The collection of HEGIS data has had an impact on the discipline and sophistication of data collection systems at institution and state levels*. This seems to be a reasonable conclusion. It was generally agreed that this discipline has facilitated the exchange of information among institutions.
>
> *The collection of HEGIS data does not impose a heavy burden on institutions since most of the data would be collected by institutions and/or states for management purposes anyway*. This conclusion seems reasonable although opponents of government regulation and data collection may argue with it. The interviewees did not see a heavy burden for ongoing systems. There is a distinct burden cost when changes are made in taxonomies, questionnaires (both of which can cause reprogramming) and/or changes in schedules.
>
> *Institutions are concerned about the uses of HEGIS for comparison purposes*. This conclusion certainly holds for comparison of

> unit costs, resource allocation, and funding. Generally institutions do not believe the data can be used for institution-to-institution comparisons because of timeliness, or lack thereof; lack of appropriate detail; differences in organization and accounting practices; and inappropriate comparisons of unlike institutions.
>
> There was general agreement that data are required from all of higher education because of differences among institutions and the uses to which the data are put. Moreover, most compilers at the institutions felt that the burden of collection would be increased rather than lessened if a sample of institutions was taken because of the increased problems in planning for and managing the collection.
>
> HEGIS data can be used for making comparisons among sectors of higher education. In fact, many would argue that it is accurate enough, handled appropriately, for making state-to-state and inter-institutional comparisons.
>
> (Andrew, 1980, pp. iv-xv)

APPENDIX B

Listing of the Nationally Recognized Accrediting Agencies and Associations

Regional Accrediting Commissions

New England Association of Schools and Colleges, Commission on Institutions of Higher Education

New England Association of Schools and Colleges, Commission on Vocational, Technical, Career Institutions

Middle States Association of Colleges and Secondary Schools, Commission on Higher Education

North Central Association of Colleges and Schools, Commission on Institutions of Higher Learning

Northwest Association of Schools and Colleges, Commission on Colleges

Southern Association of Colleges and Schools, Commission on Colleges

Western Association of Schools and Colleges, Accrediting Commission for Senior Colleges and Universities

Western Association of Schools and Colleges, Accrediting Commission for Community and Junior Colleges

National Institutional and Specialized Accrediting Bodies

National League for Nursing: associate degree program in nursing

American Association of Nurse Anesthetists, Council on Accreditation: nurse anesthesia

National Architectural Accrediting Boards: architecture

American Veterinary Medical Association: associate degree program in animal technology

American Medical Association, Committee on Allied Health Education and Accreditation: assistant to the primary care physician

APPENDIX B (cont'd)

American Library Association, Committee on Accreditation: librarianship

American Society of Landscape Architects: landscape architecture

Accrediting Bureau of Health Education Schools: medical assistant

American Medical Association, Committee on Allied Health Education and Accreditation: medical assistant

American Medical Association and Association of American Medical Colleges, Liaison Committee on Medical Education: medicine

American Medical Association and Association of American Medical Colleges, Liaison Committee on Medical Education: basic medical science

American Association for Marriage and Family Therapy Education, Commission on Accreditation for Marriage and Family Therapy: marriage and family counseling (clinical)

American Association for Marriage and Family Therapy Education, Commission on Accreditation for Marriage and Family Therapy: marriage and family counseling (graduate degree)

Accrediting Bureau of Health Education Schools: medical laboratory technician

American Medical Association, Committee on Allied Health Education and Accreditation: medical laboratory technician (associate degree)

American Medical Association, Committee on Allied Health Education and Accreditation: medical laboratory technician (certificate)

American Medical Association, Committee on Allied Health Education and Accreditation: medical record administrator

American Medical Association, Committee on Allied Health Education and Accreditation: medical record technician

American Medical Association, Committee on Allied Health Education and Accreditation: medical technologist

National Association of Schools of Music: music

American Medical Association, Committee on Allied Health Education and Accreditation: nuclear medicine technologist

National League for Nursing: baccalaureate and higher degree program in nursing

American Optometric Association, Council on Optometric Education: optometry

American Osteopathic Association, Office of Osteopathic Education: osteopathic medicine

American Medical Association, Committee on Allied Health Education and Accreditation: occupational therapy

Council on Education for Public Health: public health

American Council on Pharmaceutical Education: pharmacy

APPENDIX B (cont'd)

National Association of Schools of Art, Commission on Accreditation and Membership: art

American Speech, Language, and Hearing Association: audiology

American Medical Association, Committee on Allied Health Education and Accreditation: specialist in blood bank technology

American Association of Bible Colleges: Bible college education

American Assembly of Collegiate Schools of Business: business

Council on Education for Public Health: community health education (graduate degree)

Council on Chiropractic Education, Commission on Accreditation: chiropractic education

Council on Education for Public Health: community health, preventive medicine (graduate degree)

American Psychological Association: counseling psychology

American Medical Association, Committee on Allied Health Education and Accreditation: cytotechnology

American Dental Association, Commission on Accreditation of Dental and Dental Auxiliary Programs: dental assistant

American Dental Association, Commission on Accreditation of Dental and Dental Auxiliary Programs: dentistry

American Dental Association, Commission on Accreditation of Dental and Dental Auxiliary Programs: dental hygiene

American Dietetic Association: dietetics

American Dietetic Association: dietetic internships

American Dental Association, Commission on Accreditation of Dental and Dental Auxiliary Programs: dental laboratory technician

Engineers' Council for Professional Development: engineering

Engineers' Council for Professional Development: engineering technology

Foundation for Interior Design Education Research: interior design education; professional and technical programs

Society of American Foresters: forestry

American Board of Funeral Service Education: funeral service education

Accrediting Commission on Education for Health Services Administration: health services administration

American Medical Association, Committee on Allied Health Education and Accreditation: histologic technician

American Psychological Association: internships in clinical and counseling psychology

American Council on Education for Journalism, Accrediting Committee: journalism

Association of Independent Colleges and Schools, Accrediting Commission: private junior college of business

American Bar Association, Section of Legal Education and Admissions to the Bar: law

APPENDIX B (cont'd)

National Association for Practical Nurse Education and Service: practical nursing

National League for Nursing: practical nursing

American Podiatry Association, Council on Podiatry Education: podiatry

American Physical Therapy Association, Committee on Accreditation in Education: physical therapist

American Physical Therapy Association, Committee on Accreditation in Education: physical therapist assistant

Association of Advanced Rabbinical and Talmudic Schools, Accreditation Commission: rabbinical and Talmudic education

American Medical Association, Committee on Allied Health Education and Accreditation: radiologic technologist

American Medical Association, Committee on Allied Health Eduation and Accreditation: respiratory therapist

American Medical Association, Committee on Allied Health Education and Accreditation: respiratory therapy technician

American Psychological Association: school psychology

American Speech, Language, and Hearing Association: speech pathology

Association of Independent Colleges and Schools, Accrediting Commission: private senior college of business

American Medical Association, Committee on Allied Health Education and Accreditation: surgery technology

Council on Social Work Education, Commission on Accreditation: social work

National Council for Accreditation of Teacher Education: teacher education

Association of Theological Schools in the United States and Canada: theology

American Veterinary Medical Association: veterinary medicine

Nationally Recognized State Agency

New York State Board of Regents

APPENDIX C:

Suggestions for Improving the Wording and Format of the Fall Enrollment Survey

The wording and format of a questionnaire can be critical to the gathering of accurate data. The following are some specific suggestions for improving the wording and format of the Fall Enrollment Survey within HEGIS (Figure C-1 presents selected pages of the questionnaire).

Page 1

(1) Item 1, the block for writing the institution code number (located in the upper right hand corner), is surrounded by information for the respondent. This could be missed by respondents.

(2) Item 2, indicating the due date, is not a block in which respondents are to write. It should not have a number.

(3) Item 6 is written in all capital letters. This is much harder to read than text that is in upper and lower case. In addition, the wording of the instructions can be made clearer and shorter.

(4) Item 7 can be made more direct.

(5) The ordering of the definitions is confusing. The explanation of a multi-campus institution requires an understanding of what an institutional system is. To reduce the crowding on the pages, these definitions could be moved to the second or third pages--if they are necessary.

(6) Three different styles of type are used. This makes the form look more complicated.

Page 2

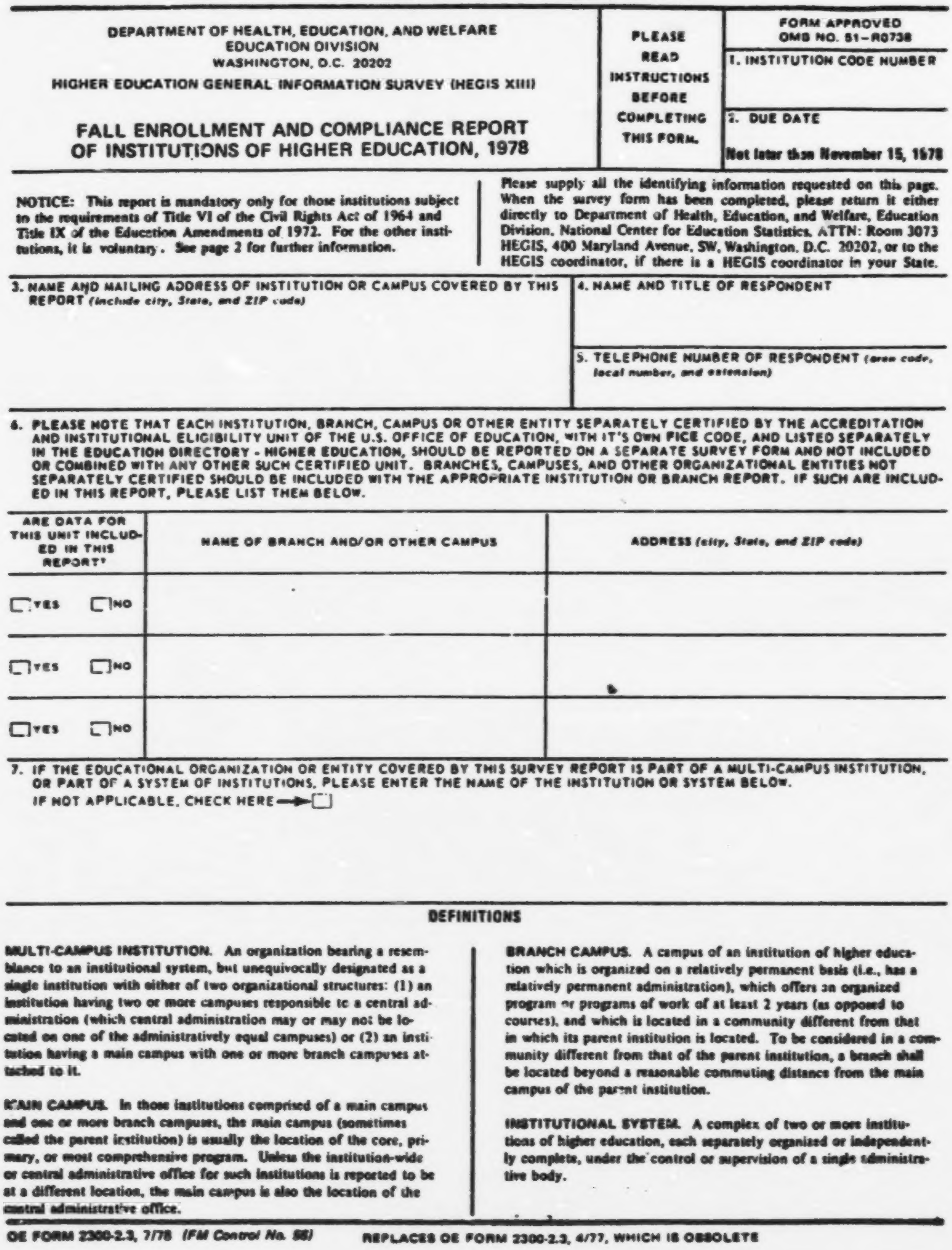
(7) A brief guide indicating that there are three parts to this report would help orient the respondent.

Page 3

(8) The purpose of the note at the top of the page is to bring the mail-in card on the last page to the respondent's attention. It should be shortened and included as part of the guide information.

Figure C-1

Reproduction of Survey Report Form



DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
EDUCATION DIVISION
WASHINGTON, D.C. 20202

HIGHER EDUCATION GENERAL INFORMATION SURVEY (HEGIS XIII)

FALL ENROLLMENT AND COMPLIANCE REPORT OF INSTITUTIONS OF HIGHER EDUCATION, 1978

PLEASE READ INSTRUCTIONS BEFORE COMPLETING THIS FORM.

FORM APPROVED
OMB NO. 51-R0738

1. INSTITUTION CODE NUMBER

2. DUE DATE
Not later than November 15, 1978

NOTICE: This report is mandatory only for those institutions subject to the requirements of Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972. For the other institutions, it is voluntary. See page 2 for further information.

Please supply all the identifying information requested on this page. When the survey form has been completed, please return it either directly to Department of Health, Education, and Welfare, Education Division, National Center for Education Statistics, ATTN: Room 3073 HEGIS, 400 Maryland Avenue, SW, Washington, D.C. 20202, or to the HEGIS coordinator, if there is a HEGIS coordinator in your State.

3. NAME AND MAILING ADDRESS OF INSTITUTION OR CAMPUS COVERED BY THIS REPORT *(include city, State, and ZIP code)*

4. NAME AND TITLE OF RESPONDENT

5. TELEPHONE NUMBER OF RESPONDENT *(area code, local number, and extension)*

6. PLEASE NOTE THAT EACH INSTITUTION, BRANCH, CAMPUS OR OTHER ENTITY SEPARATELY CERTIFIED BY THE ACCREDITATION AND INSTITUTIONAL ELIGIBILITY UNIT OF THE U.S. OFFICE OF EDUCATION, WITH IT'S OWN FICE CODE, AND LISTED SEPARATELY IN THE EDUCATION DIRECTORY - HIGHER EDUCATION, SHOULD BE REPORTED ON A SEPARATE SURVEY FORM AND NOT INCLUDED OR COMBINED WITH ANY OTHER SUCH CERTIFIED UNIT. BRANCHES, CAMPUSES, AND OTHER ORGANIZATIONAL ENTITIES NOT SEPARATELY CERTIFIED SHOULD BE INCLUDED WITH THE APPROPRIATE INSTITUTION OR BRANCH REPORT. IF SUCH ARE INCLUDED IN THIS REPORT, PLEASE LIST THEM BELOW.

ARE DATA FOR THIS UNIT INCLUDED IN THIS REPORT?	NAME OF BRANCH AND/OR OTHER CAMPUS	ADDRESS *(city, State, and ZIP code)*
☐ YES ☐ NO		
☐ YES ☐ NO		
☐ YES ☐ NO		

7. IF THE EDUCATIONAL ORGANIZATION OR ENTITY COVERED BY THIS SURVEY REPORT IS PART OF A MULTI-CAMPUS INSTITUTION, OR PART OF A SYSTEM OF INSTITUTIONS, PLEASE ENTER THE NAME OF THE INSTITUTION OR SYSTEM BELOW.
IF NOT APPLICABLE, CHECK HERE → ☐

DEFINITIONS

MULTI-CAMPUS INSTITUTION. An organization bearing a resemblance to an institutional system, but unequivocally designated as a single institution with either of two organizational structures: (1) an institution having two or more campuses responsible to a central administration (which central administration may or may not be located on one of the administratively equal campuses) or (2) an institution having a main campus with one or more branch campuses attached to it.

MAIN CAMPUS. In those institutions comprised of a main campus and one or more branch campuses, the main campus (sometimes called the parent institution) is usually the location of the core, primary, or most comprehensive program. Unless the institution-wide or central administrative office for such institutions is reported to be at a different location, the main campus is also the location of the central administrative office.

BRANCH CAMPUS. A campus of an institution of higher education which is organized on a relatively permanent basis (i.e., has a relatively permanent administration), which offers an organized program or programs of work of at least 2 years (as opposed to courses), and which is located in a community different from that in which its parent institution is located. To be considered in a community different from that of the parent institution, a branch shall be located beyond a reasonable commuting distance from the main campus of the parent institution.

INSTITUTIONAL SYSTEM. A complex of two or more institutions of higher education, each separately organized or independently complete, under the control or supervision of a single administrative body.

OE FORM 2300-2.3, 7/78 *(FM Control No. 58)*   REPLACES OE FORM 2300-2.3, 4/77, WHICH IS OBSOLETE

1

BEST COPY AVAILABLE

Figure C-1 (continued)

VOLUNTARY VERSUS REQUIRED REPORTING

This survey constitutes an integral part of the comprehensive system of statistics on higher education collected by the National Center for Education Statistics *(NCES)* as part of the Higher Education General Information Survey *(HEGIS)*.

In recent years, for the purpose of enforcing compliance with Federal regulations implementing civil rights laws applicable to institutions of higher education, the Office for Civil Rights *(OCR)* has been collecting similar enrollment data on a mandatory basis.

In order to lighten the burden on reporting institutions by eliminating the considerable duplication of effort in reporting enrollment data to two separate agencies, the two surveys have been combined into this single questionnaire and integrated into the HEGIS program.

Completion of this questionnaire is mandatory for all institutions of higher education which receive, are applicants for, or expect to be applicants for Federal financial assistance as defined in the Department of Health, Education and Welfare *(HEW)* regulation implementing Title VI *(45 CFR 80.13)*, or as defined in any HEW regulation implementing Title IX. *(See also the Instructions for completing this questionnaire.)*

Those institutions to which the regulations do not apply are not required to complete this questionnaire. However, it is hoped that institutions not subject to these provisions will voluntarily complete the entire survey--or at least the Summary page--in order that the data may represent the entire universe of higher education. As a minimum NCES requests that institutions complete columns 13 through 15 on the Summary page in order to enable NCES to continue to provide basic enrollment data serving the needs and interests of the higher education community.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

Figure C-1 (continued)

INSTRUCTIONS AND DEFINITIONS

NOTE. The card that forms the back cover of this report is for the convenience of respondents in providing preliminary data for a pre-publication release. The bottom half of the card is a self-mailer that can be detached without loosening the pages of the report.

GENERAL INSTRUCTIONS

Proofread the completed report before returning it to the National Center for Education Statistics.

This report should include only college-level students taking work creditable toward a bachelor's or higher degree or some other formal recognition below the baccalaureate.

If exact counts are lacking for a particular category of students that should be reported, include an estimate for that group.

Do NOT fill out separate forms for extension centers. Only campuses with their own FICE code numbers should be reported on separate questionnaires. Extension students should be reported on the form for the main campus.

Do NOT include in this report:

(a) Students in noncredit adult education courses.

(b) Students taking courses at home by mail, radio, or television.

(c) Students enrolled only for "short courses."

(d) Auditors.

(e) Students studying abroad if their enrollment at the reporting institution is only an administrative record and the fee is only nominal.

(f) Students in any branch campus or extension center in a foreign country.

(g) High school students taking college courses.

(h) Students known to be enrolled concurrently at another college or university, if the latter will report their enrollment *(to avoid double-counting).* Normally, the institution that will eventually grant the degree should report the student's enrollment.

NOTE. No matter what the calendar system, report on this questionnaire only those students enrolled and only those credit-hours being earned during the FALL TERM.

If you need CLARIFICATION of any item on the questionnaire that pertains to fall enrollment, please call the Survey Director, Dr. Andrew J. Pepin, NCES, (202) 245-8392, in Washington, D.C. 20202. Any questions concerning the racial/ethnic categories or major fields of study should be directed to Ms. Carol Campbell, Office for Civil Rights, (202) 245-7420, in Washington, D.C. 20201.

MAJOR FIELDS OF STUDY AND CORRESPONDING MAJOR FIELD CODES.

The listing below identifies selected categories of major fields of study, and their corresponding codes. These were taken directly from the HEGIS Taxonomy of Instructional Programs in Higher Education and aggregated into the fields listed. The field name and corresponding code number have been preprinted in the upper left-hand corner of each page. If your institution has no students enrolled in any of the designated fields, check the box as indicated. Students enrolled in the fields of Dentistry, Medicine, Veterinary Medicine and Law are not to be reported as undergraduate or graduate students but only as First-Professional students. Students in these programs requiring only 4 or 5 years beyond high school should be reported as undergraduates in the appropriate fields.

0100 – Agriculture and Natural Resources
0200 – Architecture and Environmental Design
0400 – Biological Sciences
0500 – Business and Management
0900 – Engineering
1204 – Dentistry
1206 – Medicine
1218 – Veterinary Medicine
1400 – Law
1900 – Physical Sciences
9000 – All other *(fields not included above and undecided/undeclared)*
9999 – Summary *(total enrollment).* In addition to separate reports for each of the major fields or sub-fields listed above, complete the Summary report which aggregates the enrollment data on the individual field reports.

Column 15 will be completed ONLY for the Summary report. Individual reports will NOT show data in column 15.

FILING INSTRUCTIONS – COMPLIANCE REQUIREMENTS.

Title VI of the Civil Rights Act of 1964 requires that recipients of Federal financial assistance offer their benefits and services without regard to race, color, or national origin. Title IX of the Education Amendments of 1972 requires that the benefits and services of federally assisted educational programs and activities be offered, with certain exceptions, nondiscriminatorily on the basis of sex. This report is one indicator utilized by the Office for Civil Rights in carrying out its responsibilities to verify compliance with TitleVI and Title IX. Also applicable are Section 799-A, Part H, Title VII and Section 845, Part C, Title VIII of the Public Health Service Act of 1972.

This report is to be filed by all institutions of higher education which receive, are applicants for, or expect to be applicants for Federal financial assistance as defined in the Department of Health, Education, and Welfare Regulation implementing Title VI *(45 CFR 80.13),* or as defined in any Department of Health, Education, and Welfare Regulation implementing Title IX. If your institution does not fall into any of these categories, please inform us of this fact.

Section 80.6(b) of the Regulation implementing Title VI, set forth below, and similar provisions of the Title VI Regulations of other Federal agencies, authorize collection of this information:
80.6 Compliance information
(b) Compliance reports***

> Each recipient shall keep such records and submit to the responsible Department official or his designee timely, complete and accurate compliance reports at such times, and in such form and containing such information, as the responsible Department official or his designee may determine to be necessary to enable him to ascertain whether the recipient has complied or is complying with this part. For example, recipients should have available for the Department racial and ethnic data showing the extent to which members of minority groups are beneficiaries of and participants in federally-assisted programs. In the case of any program under which a primary recipient extends Federal financial assistance to any other recipient, such other recipient shall also submit such compliance reports to the primary recipient as may be necessary to enable the primary recipient to carry out its obligations under this part.

Each institution of higher education, as well as each separately certified branch campus *(with its own FICE code number)* that is subject to the HEW civil rights regulations cited above is required to complete a separate compliance report for certain selected major fields *(listed below)* as indicated in the upper left-hand corner of each page.

RACIAL/ETHNIC CATEGORIES.

The following five racial/ethnic categories are utilized in the survey:

Black Non-Hispanic
American Indian or Alaskan Native
Asian or Pacific Islander
Hispanic
White Non-Hispanic

In addition, non-resident aliens, i.e., those members of the aforementioned groups who have not been admitted to the United States for permanent residence, should be separately identified as a sixth category; the non-resident aliens are not separately requested by racial/ethnic group, but only in totals.

The definitions for these categories are:

Non-resident alien. A person who is not a citizen of the United States and who is in this country on a temporary basis and does not

BEST COPY AVAILABLE

Figure C-1 (continued)

have the right to remain indefinitely. Resident aliens, non-citizens who have been lawfully admitted for permanent residence *(and who hold a "green card," Form I-151)*, are to be reported in the appropriate racial/ethnic categories along with United States citizens. Nonresident aliens are to be reported separately, in the columns provided, rather than in any of the five racial/ethnic categories which follow.

> **Black Non-Hispanic.** A person having origins in any of the black racial groups of Africa.
>
> **American Indian or Alaskan Native.** A person having origins in any of the original peoples of North America, and who maintains cultural identification through tribal affiliation or community recognition.
>
> **Asian or Pacific Islander.** A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands, and Samoa.
>
> **Hispanic.** A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.
>
> **White Non-Hispanic.** A person having origins in any of the original peoples of Europe, North Africa, or in the Middle East.

Racial/ethnic designations as used in this survey do not denote scientific definitions of anthropological origins. For the purpose of this report, a student may be included in the group to which he or she appears to belong, identifies with, or is regarded in the community as belonging. However, no person may be counted in more than one racial/ethnic group.

The manner of collecting the racial/ethnic information is left to the discretion of the institution provided that the system which is established results in reasonably accurate data. One acceptable method is a properly controlled system of post-enrollment self-identification by students. If a self-identification method is utilized, a verification procedure to ascertain the completeness and accuracy of student submissions should also be employed where feasible. In order to provide reasonably accurate data, the institution may require students to complete a questionnaire and/or identify themselves by name or otherwise when providing information. The fact that the information is being gathered to comply with Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972 may be disseminated in the manner and to the extent that the administration deems appropriate.

DEFINITIONS

UNDERGRADUATES. Students enrolled in a 4- or 5-year bachelor's degree program, in an associate degree program, or in a vocational or technical program that is normally terminal and results in formal recognition below the baccalaureate.

CLASS LEVEL. Whether first-year, second-year, etc., should be determined in a logical, consistent, and identifiable way. Usually, a student's class level would be based on the proportion of total requirements he has obtained toward the completion of the degree program in which he is enrolled, according to the number of years normally required to obtain them.

FIRST-TIME FRESHMEN. Entering freshmen who have not previously attended ANY college. Include students enrolled in the fall term who attended college for the first time in the summer of 1978. Also include students who entered with advanced standing *(college credits earned before graduation from high school)* at the freshman level.

OTHER FIRST-YEAR. First-year students who entered the institution before the summer of 1978.

UNCLASSIFIED STUDENTS. Not candidates for a degree or other formal award, although taking courses in regular classes with other students.

> **UNDERGRADUATE LEVEL.** Includes but is not limited to undergraduates who cannot be classified by class standing; for example, new transfer students. Also included are students who already have bachelor's degrees or awards below the baccalaureate but are taking courses at the same level or lower. Included also are "special students" and teachers taking additional undergraduate courses for certification.
>
> **POSTBACCALAUREATE LEVEL.** Includes but is not limited to "special" and other students taking first-professional or graduate courses but who are not working toward a degree.

FIRST-PROFESSIONAL STUDENTS. Students enrolled in a professional school or program which required at least 2 academic years of college work for entrance and a total of at least 6 years for a degree. Report only students in those first-professional degree programs in the field of medicine, law, and theology specified in Part A of NCES Form 2300-2.1, Degrees and Other Formal Awards Conferred. Students in programs requiring only 4 or 5 years beyond high school should be reported as undergraduates, and not in the first-professional fields.

GRADUATE STUDENTS. Students who hold the bachelor's or first-professional degree, or equivalent, and are working toward a master's or doctor's degree.

> **FIRST-YEAR.** Graduate students who have completed less than one full year of required graduate study.
>
> **BEYOND THE FIRST YEAR.** All graduate students who have completed at least one full year of graduate study toward a master's or doctor's degree.

FULL-TIME STUDENTS. Those whose academic load-coursework or other required activity is at least 75% of the normal full-time load.

FULL-TIME-EQUIVALENT (FTE) ENROLLMENT OF PART-TIME STUDENTS. Convert part-time students *(lines 15-28)* into full-time equivalents by one of the following:

> (1) Use a method already employed in your institution to compute FTE's for some other purpose.
>
> (2) Sum the credit hours for part-time students and divide by the normal full-time credit-hour load. (NORMAL FULL-TIME CREDIT-HOUR LOAD) is usually determined by dividing the total number of credits required for completing the program by the number of terms normally required to obtain them. Do not confuse this with the minimum number of credit-hours required for a student to be classified full-time (75% of a normal full-time load). NOTE: Divide by the normal, or average, full-time load, not by the minimum full-time load. For most institutions, this will be 15 credit-hours (not 12).
>
> (3) Assign a fractional value of full-time to each part-time student, appropriate to your institution, such as 1/4, 1/3, or 1/2. Remember that a student taking 3/4 *(75%)* or more of a normal full-time load should be classified as a full-time student.

BEST COPY AVAILABLE

1. NAME OF INSTITUTION	2. INSTITUTION CODE NUMBER	3. DUE DATE Not later than November 15, 1978	FORM APPROVED OMB NO. 51-R0736

99C9 – Summary *(total enrollment)*		NON-RESIDENT ALIEN		BLACK NON-HISPANIC		AMERICAN INDIAN OR ALASKAN NATIVE		ASIAN OR PACIFIC ISLANDER		HISPANIC		WHITE NON-HISPANIC		TOTAL *(sum of columns (1) through (12))*		FTE OF PART-TIME
ALL STUDENTS ENROLLED *(resident and extension)*	LINE NO.	MEN *(1)*	WOMEN *(2)*	MEN *(3)*	WOMEN *(4)*	MEN *(5)*	WOMEN *(6)*	MEN *(7)*	WOMEN *(8)*	MEN *(9)*	WOMEN *(10)*	MEN *(11)*	WOMEN *(12)*	MEN *(13)*	WOMEN *(14)*	*(15)*
I. FULL-TIME STUDENTS																
A. Undergraduates, total	01															
1. First-time freshmen	02															
2. Other first-year	03															
3. Second-year	04															
4. Third-year	05															
5. Fourth-year and beyond	06															
B. Unclassified students, total	07															
1. Undergraduate level	08															
2. Postbaccalaureate level	09															
C. First-professional students	10															
D. Graduate students, total	11															
1. First-year	12															
2. Beyond the first year	13															
TOTAL FULL-TIME STUDENTS	14															
II. PART-TIME STUDENTS																FTE OF PART-TIME
A. Undergraduates, total	15															
1. First-time freshmen	16															
2. Other first-year	17															
3. Second-year	18															
4. Third-year	19															
5. Fourth-year and beyond	20															
B. Unclassified students, total	21															
1. Undergraduate level	22															
2. Postbaccalaureate level	23															
C. First-professional students	24															
D. Graduate students, total	25															
1. First-year	26															
2. Beyond the first year	27															
TOTAL PART-TIME STUDENTS	28															
III. GRAND TOTAL ALL STUDENTS	29															

CERTIFICATION	NAME OF PERSON FURNISHING INFORMATION	TITLE	DATE
I CERTIFY that the information given above is complete, true, and correct to the best of my knowledge and belief. *(A willfully false statement is punishable by law, U.S. Code, Title 18, Section 1001.)*	SIGNATURE	TELEPHONE AREA CODE / NUMBER / EXTENSION	

OE FORM 2300-2.3, 7/78 *(FM Control No. 58)* REPLACES OE FORM 2300-2.3, 4/77, WHICH IS OBSOLETE

☆U.S. GOVERNMENT PRINTING OFFICE: 1980 624-488/1663 1-3

Figure C-1 (continued)

BEST COPY AVAILABLE

120

(9) The instruction--"Do NOT fill out separate forms for extension centers"--is explained on p. 1; it can be eliminated.

(10) The following revision of the section on Major Fields of Study and Corresponding Major Field Codes is easier to understand:

> "Below is the list of selected major fields of study and their corresponding code numbers. The list of fields is based on data from the HEGIS Taxonomy of Instructional Programs in Higher Education. There is a separate page in this report for each major field. The field name and code number is in the upper right hand corner of the page. If your institution does not have any students enrolled in a field, mark the box under the field name. Students in programs that require only 4 or 5 years beyond high school should be reported as undergraduates in their appropriate fields. Do not report students enrolled in the fields of Dentistry, Medicine, Veterinary Medicine and Law as undergraduate students. These students are First-Professional students."

(11) The section titled Filing Instructions--Compliance Requirements explains who must file and why. It interrupts the information on how to fill out the report. This should appear at the end of the instructions.

(12) This revision of the statement under the section on Racial/Ethnic Categories is clearer:

> "In addition, there is a sixth category which consists of non-resident aliens, i.e., people who have not been admitted to the United States for permanent residence. Non-resident aliens should only be included under this category; they should not be included in one of the above five racial/ethnic categories."

Page 3 and 4

(13) A statement on page 3 regarding first professional students is not completely consistent with a statement on page 4:

Page 3--under Major Fields of Study and Corresponding Major Field Codes

> "Students enrolled in the fields of Dentistry, Medicine, Veterinary Medicine and Law are not to be reported as undergraduate or graduate students but only as First-Professional students."

Page 4--under First Professional Students

> "Report only students in those first-professional degree programs in the field of medicine, law and theology as specified in Part A of NCES Form 2300-2.1, Degrees and Other Formal Awards Conferred."

Page 5

(16) The numbers in the first three boxes serve no purpose; they can be deleted.

(17) The column number has been covered up; remove the masking for this section.

These suggestions indicate some areas that may result in confusion for the respondents. A simplified version of the first two pages of the form, based on these suggestions, appears in the following pages.

Figure C-2
Example of a Revised Survey Form

FORM APPROVED
OMB No. 51-Ro738

1 Institutional Information

Higher Education General Information Survey
(HEGIS XIII)

United States Department of Education

Fall Enrollment and Compliance Report
of Institutions of Higher Education, 1981

NOTICE: This report is mandatory only for those institutions subject to the requirements of Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972. For the other institutions, it is voluntary. See pages 2 and 4 for further information.

Please read instructions before completing this form.

DUE DATE: November 15, 1981

1. Institution Code Number ______________________________

2. Name of Institution ______________________________

   Street Address ______________________________

   City ______________ State ______________ Zip Code ______________

3. Name of Respondent ______________________________

   Title ______________________________

4. Telephone Number of Respondent Area Code ______ Number ______ Extension ______

5. Below, list any branches, campuses, or other organizational entities that are not separately certified by the Accreditation and Institutional Eligibility Unit of the U.S. Department of Education. Use a separate report for each institution, branch, campus or other entity that is separately certified, has its own FICE code, and is listed separately in the Education Directory—Higher Education. Include branches, campuses, and other organizational entities that are not separately certified in the appropriate institution report.

Name of Branch and/or Other Campus	Address (City, State, and Zip Code)	Data included in this report?	
1.		☐ Yes	☐ No
2.		☐ Yes	☐ No
3.		☐ Yes	☐ No

6. Is the educational organization or entity, that is covered by this survey report, part of a multi-campus institution or part of a system of institutions?

   ☐ Yes ☐ No

   If Yes, write in the name of the institution: ______________________________

7. **When this survey form has been completed, please return before NOVEMBER 15, 1981 to:**

Department of Education
National Center for Education Statistics
Room 3073, HEGIS
400 Maryland Avenue, S.W.
Washington, D.C. 20202

OR

to the HEGIS coordinator in your State, if there is one.

OE FORM 2300-2.3, (5/81) (FM Control No. 56)

Replaces OE Form 2300-2.3, 4/77, which is obsolete

Figure C-2 (continued)

This Report consists of three parts:

		Page	Due Date
1	**Institutional Information**	1	Nov. 15, 1981
2	**Student Information**	5-16	Nov. 15, 1981
3	**Early Estimate Form**	18	Oct. 15, 1981

VOLUNTARY VERSUS REQUIRED REPORTING

This survey constitutes an integral part of the comprehensive system of statistics on higher education collected by the National Center for Education Statistics *(NCES)* as part of the Higher Education General Information Survey *(HEGIS)*.

In recent years, for the purpose of enforcing compliance with Federal regulations implementing civil rights laws applicable to institutions of higher education, the Office for Civil Rights *(OCR)* has been collecting similar enrollment data on a mandatory basis.

In order to lighten the burden on reporting institutions by eliminating the considerable duplication of effort in reporting enrollment data to two separate agencies, the two surveys have been combined into this single questionnaire and integrated into the HEGIS program.

Completion of this questionnaire is mandatory for all institutions of higher education which receive, are applicants for, or expect to be applicants for Federal financial assistance as defined in the Department of Health, Education and Welfare *(HEW)* regulation implementing Title VI *(45 CFR 80.13)*, or as defined in any HEW regulation implementing Title IX. *(See also the Instructions for completing this questionnaire.)*

Those institutions to which the regulations do not apply are not required to complete this questionnaire. However, it is hoped that institutions not subject to these provisions will voluntarily complete the entire survey–or at least the Summary page–in order that the data may represent the entire universe of higher education. As a minimum NCES requests that institutions complete columns 13 through 15 on the Summary page in order to enable NCES to continue to provide basic enrollment data serving the needs and interests of the higher education community.